GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 21:04:15 ; Search time 98.96 Seconds
(without alignments)
4735.075 Million cell updates/sec

Title: US-09-727-628-3
Perfect score: 2069
Sequence: 1 tctagatttttttttcaattc.....gggaattaaggcaaccatgg 2069

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	101.6	4.9	991	4	US-09-377-648-7	Sequence 7, Appli
	2	62.4	3.0	7218	1	US-08-232-463-14	Sequence 14, Appl
	3	41	2.0	1318	2	US-08-439-814-3	Sequence 3, Appli
	4	41	2.0	1688	2	US-08-439-814-2	Sequence 2, Appli
	5	41	2.0	2090	2	US-08-439-814-1	Sequence 1, Appli
	6	39	1.9	19124	2	US-08-487-826B-13	Sequence 13, Appl
	7	38.4	1.9	1931	4	US-09-118-442-10	Sequence 10, Appl
	8	38.4	1.9	1931	4	US-09-677-064-10	Sequence 10, Appl
	9	37.8	1.8	5622	4	US-09-067-800-3	Sequence 3, Appli
	10	37.8	1.8	5622	4	US-09-349-677-3	Sequence 3, Appli
c	11	37.4	1.8	5852	1	US-07-867-106-2	Sequence 2, Appli
c	12	37.2	1.8	602	1	US-08-764-100-8	Sequence 8, Appli
c	13	37.2	1.8	643	1	US-08-764-100-7	Sequence 7, Appli
c	14	37.2	1.8	688	4	US-08-998-416-972	Sequence 972, App
	15	37.2	1.8	740	4	US-08-998-416-971	Sequence 971, App
c	16	37.2	1.8	2993	1	US-08-764-100-2	Sequence 2, Appli
	17	37.2	1.8	2993	1	US-08-764-100-10	Sequence 10, Appl
c	18	37.2	1.8	3001	1	US-08-764-100-1	Sequence 1, Appli
c	19	36.8	1.8	701	4	US-08-998-416-701	Sequence 701, App
	20	36	1.7	2621	2	US-08-553-619B-8	Sequence 8, Appli
	21	36	1.7	7015	4	US-09-177-249-6	Sequence 6, Appli
	22	35.8	1.7	2730	2	US-08-811-897A-39	Sequence 39, Appl
	23	35.8	1.7	2730	2	US-08-855-213-39	Sequence 39, Appl
	24	35.8	1.7	2814	2	US-08-811-897A-38	Sequence 38, Appl
	25	35.8	1.7	2814	2	US-08-855-213-38	Sequence 38, Appl
	26	35.8	1.7	6769	1	US-08-480-784-20	Sequence 20, Appl
	27	35.8	1.7	6769	1	US-08-483-553-20	Sequence 20, Appl
	28	35.8	1.7	6769	1	US-08-487-002-20	Sequence 20, Appl
	29	35.8	1.7	6769	1	US-08-483-554B-20	Sequence 20, Appl
	30	35.8	1.7	6769	1	US-08-488-011B-20	Sequence 20, Appl
	31	35.8	1.7	6769	4	US-08-850-727-20	Sequence 20, Appl
	32	35.8	1.7	6769	5	PCT-US95-10202-20	Sequence 20, Appl
	33	35.8	1.7	6769	5	PCT-US95-10203-20	Sequence 20, Appl
	34	35.8	1.7	6769	5	PCT-US95-10220-20	Sequence 20, Appl
	35	35.6	1.7	1826	4	US-09-286-691-11	Sequence 11, Appl
	36	35.6	1.7	1826	4	US-09-687-147-11	Sequence 11, Appl
	37	35.6	1.7	2450	2	US-08-463-081B-5	Sequence 5, Appli
	38	35.6	1.7	2450	2	US-08-461-379A-5	Sequence 5, Appli
	39	35.6	1.7	2450	2	US-08-462-390B-5	Sequence 5, Appli
	40	35.6	1.7	2450	3	US-08-463-074B-5	Sequence 5, Appli
	41	35.6	1.7	2450	3	US-08-465-585C-5	Sequence 5, Appli
	42	35.6	1.7	2450	3	US-08-652-446-5	Sequence 5, Appli
c	43	35.6	1.7	5613	2	US-08-463-418-1	Sequence 1, Appli
	44	35.4	1.7	2287	4	US-08-845-258-8	Sequence 8, Appli
	45	35.4	1.7	2287	4	US-08-723-142A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-377-648-7
; Sequence 7, Application US/09377648
; Patent No. 6225529
; GENERAL INFORMATION:
; APPLICANT: Lappegard, Kathryn
; APPLICANT: Martino-Catt, Susan
; TITLE OF INVENTION: Seed-Preferred Promoters
; FILE REFERENCE: 0869
; CURRENT APPLICATION NUMBER: US/09/377,648
; CURRENT FILING DATE: 1999-08-19
; EARLIER APPLICATION NUMBER: US 60/097,233
; EARLIER FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 991
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)...(991)
US-09-377-648-7

Query Match 4.9%; Score 101.6; DB 4; Length 991;
Best Local Similarity 68.9%; Pred. No. 2.5e-20;
Matches 230; Conservative 0; Mismatches 64; Indels 40; Gaps 5;

```
Qy   1696 ctgccgacggccgacggcctcccatttcgctcccctcctactcctatgctgcggtccagc 1755
          | |||| | ||||||| ||||||||||||||||||     |||||  ||||||||| |||
Db    659 cggccgccagccgacgtcctcccatttcgctcccc-----ctcctccgctgcggtcgagc 713

Qy   1756 aaaagttcgggccttccggcaatccgccggcgcccgtcggctcaaatcgcat-------- 1807
          |||||||| |||| |||||||||| ||||| |||| ||| ||||||||  ||
Db    714 aaaagttccggccatccggcaatccccccggcacccggcggttcaaatcgtatctttctga 773

Qy   1808 -ctaccgcggctagaagctctctcttcc--tccctccgatccggtggggtccatttcctt 1864
           ||  |||||||  |  || ||| |||  ||||| ||||||||||||||| |||| ||
Db    774 cctgacgcggctacgacgtcgctcctccggtcccttcgatccggtggggtccgtttcttt 833

Qy   1865 caattgtggcagtggccgtc------------------tcgaaccctctataaatccccc 1906
          |||  | |||   |  | |                  ||||||||||||||||||||
Db    834 caagcgcggcctcgctggccgcctcgtggcagtgaccgtcgaaccctctataaatcccgt 893

Qy   1907 accccggacacccttccccgaccacacagcccaacaacaaggagcgcggcggccccctcct 1966
          |||||  |||||||||| ||| |||||| ||      |||||    |   |   ||||||
Db    894 gccccgagcacccttcctcgatcacacaacc------caaggcagccacagcagcctcct 947

Qy   1967 tccttcctcccacttctctcgcgcggcgctcgct 2000
```

```
          |||| |   |    || |||| || |||||||||
Db    948 tcctcctctcactctcgctcgtgctgcgctcgct 981

RESULT   2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             3.0%;  Score 62.4;  DB 1;  Length 7218;
  Best Local Similarity   0.8%;  Pred. No. 5.3e-08;
  Matches    3;  Conservative  234;  Mismatches  135;  Indels    0;  Gaps    0;

Qy   1652 ggtatcctatccattgtcatcgctcacagtcacagaggtagccactgccgacggccgacg 1711
          | | :::: ::: :: :: :: ::: :  :: :     :  :: :: ::  :  ::  :
Db   1064 GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123

Qy   1712 gcctcccatttcgctcccctcctactcctatgctgcggtccagcaaaagttcgggccttc 1771
          :::::: :::: :::::::::: ::::: : :: :  ::: :     :::   :::::
Db   1124 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1183

Qy   1772 cggcaatccgccggcgcccgtcggctcaaatcgcatctaccgcggctagaagctctctct 1831
          :  :  ::: :: : ::: ::  :::   :: : ::: :: :  ::     ::::::::
Db   1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243

Qy   1832 tcctccctccgatccggtggggtccatttccttcaattgtggcagtggccgtctcgaacc 1891
          :::::::::  :::  :    ::: :::::::  :: :  :  :  :: :::   ::
Db   1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303

Qy   1892 ctctataaatcccccaccccggacaccccttccccgaccacacagcccaacaacaaggagc 1951
          :::: :   :::::: ::::   : :::::::::  :: : :  :::  :  :     :
Db   1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363

Qy   1952 gcggcggcccctccttccttcctcccacttctctcgcgcggcgctcgcttacctcgcctc 2011
          :  : ::::::::::::::::::: :::::::: : :  : ::: ::: :::: ::::
Db   1364 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1423

Qy   2012 gcattccgttcg 2023
          : :::: :::|
Db   1424 YYYYYYYYYYYG 1435


RESULT   3
US-08-439-814-3
; Sequence 3, Application US/08439814
; Patent No. 5968735
;  GENERAL INFORMATION:
;    APPLICANT:  STEIN, Ulrike
;    APPLICANT:  WALTHER, Wolfgang
;    TITLE OF INVENTION:  VECTOR FOR THE EXPRESSION OF
;    TITLE OF INVENTION:  THERAPY-RELEVANT GENES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,814
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, Sharon N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  P1614-5015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B (EPO)
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-439-814-3
```

```
  Query Match             2.0%;  Score 41;  DB 2;  Length 1318;
  Best Local Similarity   63.9%;  Pred. No. 0.038;
  Matches   62;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy    702 atgttaaacatttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaa 761
          || |||||  || | ||||||||| |   | | ||||  |||  | | |     |||| |
Db    462 ATATTAAATGTTGGCAGTAAATATGGAAGGAAATTACAACTAATGTAATATGCTAAAACA 521

Qy    762 tgtagtattgttttatataatttaataaaactgtaa 798
          ||   | ||  |||   |||||| ||| ||| |||||
Db    522 TGCTATGTTTATTTTACTAATTTGAATTAAAATGTAA 558


RESULT   4
US-08-439-814-2
; Sequence 2, Application US/08439814
; Patent No. 5968735
;  GENERAL INFORMATION:
;    APPLICANT:  STEIN, Ulrike
;    APPLICANT:  WALTHER, Wolfgang
;    TITLE OF INVENTION:  VECTOR FOR THE EXPRESSION OF
;    TITLE OF INVENTION:  THERAPY-RELEVANT GENES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,814
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, Sharon N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  P1614-5015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B (EPO)
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1688 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
```

US-08-439-814-2

```
  Query Match             2.0%;  Score 41;  DB 2;  Length 1688;
  Best Local Similarity   63.9%;  Pred. No. 0.046;
  Matches   62;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy    702 atgttaaacatttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaa 761
          || |||||  || | ||||||||| |   | | ||||  |||  | | |     |||| |
Db    832 ATATTAAATGTTGGCAGTAAATATGGAAGGAAATTACAACTAATGTAATATGCTAAAACA 891

Qy    762 tgtagtattgttttatataatttaataaaactgtaa 798
          ||   | ||  |||   |||||| ||| ||| |||||
Db    892 TGCTATGTTTATTTTACTAATTTGAATTAAAATGTAA 928


RESULT   5
US-08-439-814-1
; Sequence 1, Application US/08439814
; Patent No. 5968735
;  GENERAL INFORMATION:
;    APPLICANT:  STEIN, Ulrike
;    APPLICANT:  WALTHER, Wolfgang
;    TITLE OF INVENTION:  VECTOR FOR THE EXPRESSION OF
;    TITLE OF INVENTION:  THERAPY-RELEVANT GENES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,814
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, Sharon N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  P1614-5015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B (EPO)
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4238778.7
;      FILING DATE:  12-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/DE93/01086
;      FILING DATE:  10-NOV-1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2090 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  DNA (genomic)
US-08-439-814-1


  Query Match             2.0%;  Score 41;  DB 2;  Length 2090;
  Best Local Similarity   63.9%;  Pred. No. 0.053;
  Matches   62;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy    702 atgttaaacatttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaa 761
          || ||||| || | |||||||||| |   | | ||||  ||| | | |     |||| |
Db   1234 ATATTAAATGTTGGCAGTAAATATGGAAGGAAATTACAACTAATGTAATATGCTAAAACA 1293

Qy    762 tgtagtattgttttatataatttaataaaactgtaa 798
          ||   | ||  |||   |||||| ||| ||| |||||
Db   1294 TGCTATGTTTATTTTACTAATTTGAATTAAAATGTAA 1330


RESULT   6
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhaun
;    APPLICANT:  Wellems, Thomas E.
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19124 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-487-826B-13


  Query Match             1.9%;  Score 39;  DB 2;  Length 19124;
  Best Local Similarity   46.5%;  Pred. No. 1;
  Matches  126;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy    529 taaaagcagagatgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgct 588
          |||||  | |  | || ||  |     |   | |   |||    ||||| ||||   |
Db  15604 TAAAAAAAAATTTTATATAAAAAAAAAATGATTATAAAAAAAATAAAAACAAAAGAAGAA 15663

Qy    589 ccacaataacgacaaaacaaagttgtattgtatggcctaaattacagcacactgacacca 648
            | || |||   |||| |||       | |||     |||  | |  | |   | |  |
Db  15664 AAAAAAAAACATTAAAAAAAAAAAAAATATATATCATAAAAACAAAAAAAAAAGAAAAAAA 15723

Qy    649 cacgtatattattctctctccattatcacaggatgtaactgtaaaaattttgtatgttaa 708
          |  |  |         | |  || || | | |   ||   ||||||  |  ||    ||
Db  15724 TATATTAAAATAAAAATATATATCATAAAATAAAAAAAAATTAAAAAAATGTTAAAAAAA 15783

Qy    709 acatttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaatgtagta 768
          | || | ||   || ||     |  |||||  |  |   ||   | | | |||    |  |
Db  15784 AAATATATACATAAAATAAAAAAAATTTATTTAAATAAAAAAAAATAATAAATAAAAAAA 15843

Qy    769 ttgttttatataattttaataaaactgtaaa 799
          ||   ||| ||||    ||  ||    ||||
Db  15844 TTTAATTAAATAAAAAAAAATAATAAATAAA 15874


RESULT   7
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
;  APPLICANT: Martino-Catt, Susan J.
;  APPLICANT: Wang, Hongyu
;  APPLICANT: Beach, Larry R.
;  APPLICANT: Wang, Xun
;  APPLICANT: Bowen, Benjamin A.
;  TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
;  TITLE OF INVENTION: Plants and Uses Thereof
;  FILE REFERENCE: 0706
;  CURRENT APPLICATION NUMBER: US/09/118,442B
;  CURRENT FILING DATE: 1998-07-17
;  EARLIER APPLICATION NUMBER: 60/055,446
;  EARLIER FILING DATE: 1997-08-11
;  EARLIER APPLICATION NUMBER: 60/055,526
;  EARLIER FILING DATE: 1997-08-08
;  EARLIER APPLICATION NUMBER: 60/053,944
;  EARLIER FILING DATE: 1997-07-28
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1931
;   TYPE: DNA
;   ORGANISM: Zea mays
US-09-118-442-10


  Query Match             1.9%;  Score 38.4;  DB 4;  Length 1931;
  Best Local Similarity   70.7%;  Pred. No. 0.3;
  Matches   65;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy   1947 ggagcgcggcggcccctccttccttcctcccacttctctcgcgcggcgctcgcttacctc 2006
           ||  || |  |  |||||||||| |   |    || ||||||| |||||||| ||||||
Db      1 ggcacgagcagcagcctccttcctcctctcactctcgctcgcgctgcgctcgc-tacctc 59

Qy   2007 gcctcgcattccgttcgagcaggggagcggca 2038
          || ||||||||| |||||  || || | || |
Db     60 gcttcgcattccattcgaaaagaggggaggaa 91


RESULT   8
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
;  APPLICANT:  Martino-Catt, Susan J.
;  APPLICANT:  Wang, Hongyu
;  APPLICANT:  Beach, Larry R.
;  TITLE OF INVENTION:  Genes Controlling Phytate Metabolism in
;  TITLE OF INVENTION:  Plants and Uses Thereof
;  FILE REFERENCE:  0706D
;  CURRENT APPLICATION NUMBER: US/09/677,064
```

```
;  CURRENT FILING DATE:  2000-09-29
;  PRIOR APPLICATION NUMBER:  60/055,446
;  PRIOR FILING DATE:  1997-08-11
;  PRIOR APPLICATION NUMBER:  60/055,526
;  PRIOR FILING DATE:  1997-08-08
;  PRIOR APPLICATION NUMBER:  60/053,944
;  PRIOR FILING DATE:  1997-07-28
;  PRIOR APPLICATION NUMBER:  09/118,442
;  PRIOR FILING DATE:  1998-07-17
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1931
;   TYPE: DNA
;   ORGANISM: Zea mays
US-09-677-064-10


  Query Match             1.9%;  Score 38.4;  DB 4;  Length 1931;
  Best Local Similarity   70.7%;  Pred. No. 0.3;
  Matches   65;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy   1947 ggagcgcggcggccccctccttccttcctcccacttctctcgcgcggcgctcgcttacctc 2006
           ||  || |  |   |||||||||| |   |    || ||||||| |||||||| ||||||
Db      1 ggcacgagcagcagcctccttcctcctctcactctcgctcgcgctgcgctcgc-tacctc 59

Qy   2007 gcctcgcattccgttcgagcaggggagcggca 2038
           || |||||||| |||||  || || | || |
Db     60 gcttcgcattccattcgaaaagaggggaggaa 91


RESULT   9
US-09-067-800-3
; Sequence 3, Application US/09067800
; Patent No. 6198024
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Ferrandiz, Cristina
;    TITLE OF INVENTION:  Seed Plants Characterized by Delayed
;    TITLE OF INVENTION:  Seed Dispersal
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/067,800
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2948
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5622 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..5622
;      OTHER INFORMATION:  /label= AGL1_promoter
;      OTHER INFORMATION:  /note= "Nucleotide sequence of the AGL1 promoter."
US-09-067-800-3


  Query Match             1.8%;  Score 37.8;  DB 4;  Length 5622;
  Best Local Similarity   50.8%;  Pred. No. 0.97;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy    114 aaaaaaaacctatttataacaaacatctgccaagaatacaattctttttatacacaacttta 173
          ||||| ||   ||||| || |||  | |         |   |||||| | ||  ||| | |
Db   2271 AAAAATAAGAGATTTAAAATAAATGTTTCTTCTTTCTCTGATTCTTGTGTAACCAATTCA 2330

Qy    174 tatgtgagttctttttctcttgtaactcttattaataaaacatttttggctattaaataa 233
          |  ||  | | | |||||    ||   ||||| || |||    |||  |  |  ||
Db   2331 TGGGTTTGATATGTTTCTTGGTTATTGCTTATCAACAAAGAGATTTGATCATTATAAAGT 2390

Qy    234 tggcaactaagttagcaccactgtaattagattttgtctggaacaatttctctgact 290
           |   | ||| |    | |||   || |   ||| | | |   |  || || | |
Db   2391 AGATTAATAACTCTTAAACACACAAAGTTTCTTTATTTTTTAGTTACATCCCTAATT 2447


RESULT  10
US-09-349-677-3
; Sequence 3, Application US/09349677
; Patent No. 6288305
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Ferrandiz, Cristina
;    TITLE OF INVENTION:  Seed Plants Characterized by Delayed
;    TITLE OF INVENTION:  Seed Dispersal
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/349,677
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/067,800
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2948
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5622 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..5622
;      OTHER INFORMATION:  /label= AGL1_promoter
;      OTHER INFORMATION:  /note= "Nucleotide sequence of the AGL1 promoter."
US-09-349-677-3


  Query Match             1.8%;  Score 37.8;  DB 4;  Length 5622;
```

```
  Best Local Similarity   50.8%;  Pred. No. 0.97;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy     114 aaaaaaaacctatttataacaaacatctgccaagaatacaattctttttatacacaactta 173
           ||||| ||   ||||| || |||  | |         |   |||||| | ||  ||| | |
Db    2271 AAAAATAAGAGATTTAAAATAAATGTTTCTTCTTTCTCTGATTCTTGTGTAACCAATTCA 2330

Qy     174 tatgtgagttctttttctcttgtaactcttattaataaaacatttttggctattaaataa 233
           |  ||  | | | |||||    ||   ||||| || |||    |||   |  |  ||
Db    2331 TGGGTTTGATATGTTTCTTGGTTATTGCTTATCAACAAAGAGATTTGATCATTATAAAGT 2390

Qy     234 tggcaactaagttagcaccactgtaattagattttgtctggaacaatttctctgact 290
           |   | ||| |    | |||   || |   |||  | |  |   |  || || | |
Db    2391 AGATTAATAACTCTTAAACACACAAAGTTTCTTTATTTTTTAGTTACATCCCTAATT 2447


RESULT  11
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
;  GENERAL INFORMATION:
;    APPLICANT:  Slade, Martin B
;    APPLICANT:  Chang, Andy C M
;    APPLICANT:  Williams, Keith L
;    TITLE OF INVENTION:  Improved Plasmid Vectors for Cellular
;    TITLE OF INVENTION:  Slime Moulds of the Genus Dictyostelium
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
;      STREET:  One Liberty Place 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/867,106
;      FILING DATE:  19920625
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ 7187
;      APPLICATION NUMBER:  PCT/AU90/00530
;      FILING DATE:  02-NOV-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feeney, Joanne Longo
;      REGISTRATION NUMBER:  35,134
;      REFERENCE/DOCKET NUMBER:  RICE-0002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5852 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
US-07-867-106-2


  Query Match           1.8%;  Score 37.4;  DB 1;  Length 5852;
  Best Local Similarity   46.6%;  Pred. No. 1.3;
  Matches  156;  Conservative    0;  Mismatches  176;  Indels    3;  Gaps    1;

Qy       6 attttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggagaaatt 65
           | ||||||  | ||    |  | | || |   | |   ||| | ||| | |    ||||
Db    1952 AGTTTTTTATAGTTTTTGCATATTTAAAAATAACTTTTAATTTTAAATTGATTTTTAATT 1893

Qy      66 atatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaaccta 125
           ||  |  |  || | | ||  || |    ||| |       |  |||||||||   |
Db    1892 ATGAGATCTAATAAAAAAAAAAAATTTTAAAATTTAAAAAAAAAAGAAAAAAAAAAAAAA 1833

Qy     126 tttataacaaacatctgccaagaatacaattc---ttttatacacaacttatatgtgagt 182
             || ||   |          | ||||  ||||    || ||| |  |  ||||| |   |
Db    1832 AGTAGAATTTATTAAAAATTAAAATATTATTCAATCTTAATAAATTAAGTATATATCGAT 1773

Qy     183 tctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaacta 242
               | | || | |    |||  | |||| |  |  |   | | |||     ||| ||
Db    1772 AGGCAATTTATTTTTATATCTATCTAAAAAAAAACTAGGAAAAATGAATGTCATCAAATA 1713

Qy     243 agttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctattt 302
             |   | || | |  ||  |||| |    || || |    |||| || ||   |
Db    1712 GTATTTTAACATTTTTTTTTTTTTTTTTAAAAAAAAAGTGTCATGACAAAAAAAAAAAAG 1653

Qy     303 ggactgtccttttgccaaacaagtagaaaatggaa 337
           | |    |       ||| ||  | | |  ||||
Db    1652 TGTCATGACAAAAAAAAAAAAAAAAAAGAGGGGAA 1618


RESULT  12
US-08-764-100-8/c
; Sequence 8, Application US/08764100
; Patent No. 5773700
;  GENERAL INFORMATION:
;    APPLICANT:  van Grinsven J., Martinus Q.
;    APPLICANT:  De Haan, Petrus T.
;    APPLICANT:  Gielen L., Johannes J.
;    APPLICANT:  Peters, Dirk
;    APPLICANT:  Goldbach, Robert W.
;    TITLE OF INVENTION:  Improvements in or Relating to Organic
;    TITLE OF INVENTION:  Compounds
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,100
;      FILING DATE:  06-DEC-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/214,064
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/032,235
;      FILING DATE:  17-MAR-1993
;      APPLICATION NUMBER:  GB 9206016.9
;      FILING DATE:  19-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5773700ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:  137-1061
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 354-3592
;      TELEFAX:  (415) 857-1125
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  602 base pairs
;      TYPE:  nucleic acid
```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-764-100-8

Query Match 1.8%; Score 37.2; DB 1; Length 602;
Best Local Similarity 49.0%; Pred. No. 0.3;
Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy    113 aaaaaaaaacctatttataacaaacatctgccaagaatacaattcttttatacacaactt 172
          |||| |||| | |  | || ||||    | || |  |  |    ||| |   ||| ||
Db    312 AAAACAAAAACAAAAAACAAAAAACAAAAACAAAAACAAATACAAATTTTTGGCCAATTT 253

Qy    173 atatgtgagttctttttctcttgtaactcttattaataaaacatttttggctattaaata 232
          |  ||   ||  |||    | ||    ||| ||||| |    ||||      | | ||
Db    252 GTGAGTAGATTGGATTTAAAGTTTATAATTTACTAATACATTCATTTTAAAGCATTATTA 193

Qy    233 atggcaactaagttagcaccactgtaattagattttgtctggaacaatttctctgactaa 292
          | |  | | || |  || | |  |  ||    |||||  |   |    || | ||| |
Db    192 AAGCAACCAAATTGTGCCCGAAGGCCATCCTTTTTTGGTTTTTATTTATTTTATGATTTT 133

Qy    293 gaagctatttggactgtccttt 314
              |||||    | |  |||
Db    132 TTTTGTATTTTTTATATATTTT 111
```

RESULT 13
US-08-764-100-7/c
; Sequence 7, Application US/08764100
; Patent No. 5773700
; GENERAL INFORMATION:
; APPLICANT: van Grinsven J., Martinus Q.
; APPLICANT: De Haan, Petrus T.
; APPLICANT: Gielen L., Johannes J.
; APPLICANT: Peters, Dirk
; APPLICANT: Goldbach, Robert W.
; TITLE OF INVENTION: Improvements in or Relating to Organic
; TITLE OF INVENTION: Compounds
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,100
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,064
; FILING DATE:
; APPLICATION NUMBER: US 08/032,235
; FILING DATE: 17-MAR-1993
; APPLICATION NUMBER: GB 9206016.9
; FILING DATE: 19-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5773700ris, Allen E.
; REGISTRATION NUMBER: 34,490
; REFERENCE/DOCKET NUMBER: 137-1061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 354-3592
; TELEFAX: (415) 857-1125
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-764-100-7

Query Match 1.8%; Score 37.2; DB 1; Length 643;
Best Local Similarity 49.0%; Pred. No. 0.31;
Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy    113 aaaaaaaaacctatttataacaaacatctgccaagaatacaattcttttatacacaactt 172
          |||| |||| | |  | || ||||    | || |  |  |    ||| |   ||| ||
Db    315 AAAACAAAAACAAAAAACAAAAAACAAAAACAAAAACAAATACAAATTTTTGGCCAATTT 256

Qy    173 atatgtgagttctttttctcttgtaactcttattaataaaacatttttggctattaaata 232
          |  ||   ||  |||    | ||    ||| ||||| |    ||||      | | ||
Db    255 GTGAGTAGATTGGATTTAAAGTTTATAATTTACTAATACATTCATTTTAAAGCATTATTA 196

Qy    233 atggcaactaagttagcaccactgtaattagattttgtctggaacaatttctctgactaa 292
          | |  | | || |  || | |  |  ||    |||||  |   |    || | ||| |
Db    195 AAGCAACCAAATTGTGCCCGAAGGCCATCCTTTTTTGGTTTTTATTTATTTTATGATTTT 136

Qy    293 gaagctatttggactgtccttt 314
              |||||    | |  |||
Db    135 TTTTGTATTTTTTATATATTTT 114
```

RESULT 14
US-08-998-416-972/c
; Sequence 972, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 972:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1598UP
US-08-998-416-972


  Query Match             1.8%;  Score 37.2;  DB 4;  Length 688;
  Best Local Similarity   48.2%;  Pred. No. 0.33;
  Matches  105;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy      42 acaatctaaaaatcaggagaaattatatggccatattatagaagcaactaaataaaatgt 101
           | ||  |||  || |    || ||| ||    || | ||       |    ||| | | |
Db     443 ATAACATAACTATTAACTAAATTTAAATATAAATTTGATTTTCTATATGCCATATATTTT 384

Qy     102 gcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaatacaattctttt 161
             |  |   |  |  |||    | | ||| ||| ||  |   | |||| ||||||||
Db     383 ATTATTAAAATATTATTAAATTATTATTTAATAAATATTAGATTATAATATAATTCTTTA 324

Qy     162 atacacaacttatatgtgagttctttttctcttgtaactcttattaataaaacatttttg 221
           | | |   |||   | | ||  | |  |  | ||    |||||| | |   || |||
Db     323 TAATAAATTATATTATTTAATTAATATATTAATTTATTATTTATTATTTATTAATATTTA 264

Qy     222 gctattaaataatggcaactaagttagcaccactgtaa 259
             || |   ||  || |   ||  ||  |   |  |||
Db     263 TATAATCTTTATAGGAATTGAACCTAATAAACCATTAA 226


RESULT  15
US-08-998-416-971
; Sequence 971, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  971:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  740 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1598RP
US-08-998-416-971


  Query Match             1.8%;  Score 37.2;  DB 4;  Length 740;
  Best Local Similarity   49.0%;  Pred. No. 0.35;
  Matches   99;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy      42 acaatctaaaaatcaggagaaattatatggccatattatagaagcaactaaataaaatgt 101
           | ||  |||  || |    || ||| ||    || | ||       |    ||| | | |
Db     451 ATAACATAACTATTAACTAAATTTAAATATAAATTTGATTTTCTATATGCCATATATTTT 510

Qy     102 gcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaatacaattctttt 161
             |  |   |  |  |||    | | ||| ||| ||  |   | |||| ||||||||
Db     511 ATTATTAAAATATTATTAAATTATTATTTAATAAATATTAGATTATAATATAATTCTTTA 570

Qy     162 atacacaacttatatgtgagttctttttctcttgtaactcttattaataaaacatttttg 221
           | | |   |||   | | ||  | |  |  | ||    |||||| | |   || |||
Db     571 TAATAAATTATATTATTTAATTAATATATTAATTTATTATTTATTATTTATTAATATTTA 630

Qy     222 gctattaaataatggcaactaa 243
             || |   ||  || || | |
Db     631 TATAATCTTTATAGGGAATTGA 652


Search completed: December 19, 2001, 22:39:44
Job time: 5729 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 19:59:19 ; Search time 2941.36 Seconds
(without alignments)
11604.382 Million cell updates/sec

Title: US-09-727-628-3
Perfect score: 2069
Sequence: 1 tctagatttttttcaattc.....gggaattaaggcaaccatgg 2069

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
1		2069	100.0	2069	6	AX151209	AX151209 Sequence
2		119	5.8	1959	6	AX151207	AX151207 Sequence
3		105.2	5.1	753	6	AX090098	AX090098 Sequence
4		101.6	4.9	991	6	AR148903	AR148903 Sequence
5	c	62.6	3.0	14867	3	AE001398	AE001398 Plasmodiu
6		62.4	3.0	7218	6	I66494	I66494 Sequence 14
7		61	2.9	1141	6	AX083744	AX083744 Sequence
8		60	2.9	175271	9	AL354718	AL354718 Human DNA
9	c	59.2	2.9	128769	2	AL139261	AL139261 Homo sapi
10	c	55	2.7	1669	3	AB035880	AB035880 Chaitoreg
11	c	54.2	2.6	149733	2	AC087568	AC087568 Pan trogl
12	c	53.4	2.6	171769	8	F16N3	AC007519 Sequence
13	c	53	2.6	1141	6	AX083744	AX083744 Sequence
14		51.6	2.5	192929	2	AC005505	AC005505 Plasmodiu
15		51.2	2.5	104992	2	AC005504	AC005504 Plasmodiu
16		51.2	2.5	169546	2	AC004157	AC004157 Plasmodiu
17	c	51.2	2.5	318221	2	PFMAL13P3	AL049184 Plasmodiu
18		50.4	2.4	2177	3	XV18SRRN	X74763 X.vesparum
19		50.4	2.4	3316	3	XV18RRNA	X77784 X.vesparum
20	c	50.4	2.4	153477	2	AC006278	AC006278 Plasmodiu
21	c	50.4	2.4	203012	2	AC010744	AC010744 Homo sapi
22	c	50.4	2.4	205429	2	AC005506	AC005506 Plasmodiu
23		50.2	2.4	163443	2	AC006280	AC006280 Plasmodiu
24		50.2	2.4	183584	9	AC012492	AC012492 Homo sapi
25	c	49.8	2.4	165957	2	AC073951	AC073951 Homo sapi
26		49.8	2.4	176639	9	AC016987	AC016987 Homo sapi
27		49.6	2.4	130282	2	AC078973	AC078973 Homo sapi
28	c	49.4	2.4	137339	2	AC087318	AC087318 Homo sapi
29	c	49.2	2.4	43387	8	ATK10A8	AL391151 Arabidops
30	c	49.2	2.4	155711	3	VYIVD10	AL360354 Plasmodiu
31	c	49.2	2.4	170357	2	AL359551	AL359551 Homo sapi
32		49	2.4	140414	2	AF377947	AF377947 Oryza sat
33	c	48.6	2.3	183584	9	AC012492	AC012492 Homo sapi
34	c	48.4	2.3	165756	2	AL355138	AL355138 Homo sapi
35		48.2	2.3	85127	2	AC025978	AC025978 Homo sapi
36	c	48.2	2.3	178500	2	AC011981	AC011981 Homo sapi
37	c	48	2.3	104992	2	AC005504	AC005504 Plasmodiu
38		48	2.3	113254	9	AL137009	AL137009 Human DNA
39	c	48	2.3	169546	2	AC004157	AC004157 Plasmodiu
40		48	2.3	179260	2	AL591479	AL591479 Homo sapi
41		48	2.3	236680	2	AC024120	AC024120 Homo sapi
42	c	47.8	2.3	742	3	CPU69572	U69572 Mitochondri
43		47.8	2.3	12573	3	PFSC03006	Z97349 Plasmodium
44	c	47.8	2.3	20598	6	AX067857	AX067857 Sequence
45	c	47.8	2.3	81631	9	HS745I14	AL033532 Human DNA

ALIGNMENTS

RESULT 1
AX151209
LOCUS AX151209 2069 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 3 from Patent WO0140440.
ACCESSION AX151209
VERSION AX151209.1 GI:14533377
KEYWORDS .
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2069)
AUTHORS Armstrong,K., Hey,T.D., Folkerts,O., Smith,K.A. and Hopkins,N.L.
TITLE Maize mip synthase promoter
JOURNAL Patent: WO 0140440-A 3 07-JUN-2001;
Dow AgroSciences LLC (US)
FEATURES Location/Qualifiers
source 1. .2069
/organism="Zea mays"

/db_xref="taxon:4577"
BASE COUNT 567 a 531 c 435 g 536 t
ORIGIN

Query Match 100.0%; Score 2069; DB 6; Length 2069;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2069; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 tctagatttttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggag 60
Db       1 TCTAGATTTTTTTTCAATTCACCCCGAGTAAATATCCAATCACAATCTAAAAATCAGGAG 60

Qy      61 aaattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaa 120
Db      61 AAATTATATGGCCATATTATAGAAGCAACTAAATAAAATGTGCGTTGTATTGAAAAAAAA 120

Qy     121 acctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtga 180
Db     121 ACCTATTTATAACAAACATCTGCCAAGAATACAATTCTTTTATACACAACTTATATGTGA 180

Qy     181 gttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaac 240
Db     181 GTTCTTTTTCTCTTGTAACTCTTATTAATAAAACATTTTTGGCTATTAAATAATGGCAAC 240

Qy     241 taagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctat 300
Db     241 TAAGTTAGCACCACTGTAATTAGATTTTGTCTGGAACAATTTCTCTGACTAAGAAGCTAT 300

Qy     301 ttggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctc 360
Db     301 TTGGACTGTCCTTTTGCCAAACAAGTAGAAAATGGAACCGCTCCTTAAAAAACCATTCTC 360

Qy     361 acatcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgc 420
Db     361 ACATCGCTGGGTGCTGAATAAAACTGAAAACATTAGCTTTTTATAGCTCTCGCTCTCTGC 420

Qy     421 tagtatgtgttataaaatcattttaccaattaccttttttaaataactgtacgtagtttca 480
Db     421 TAGTATGTGTTATAAAATCATTTTACCAATTACCTTTTTAAATAACTGTACGTAGTTTCA 480

Qy     481 tcagtagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagaga 540
Db     481 TCAGTAGAACTACTCACGGAGCTAAAACAAAAAAAGTTGTTCTACTGATAAAAGCAGAGA 540

Qy     541 tgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacga 600
Db     541 TGATGTATGACCGTGACCGTGAGCTAAAGTCCAAAAAAAAAAACTGCTCCACAATAACGA 600

Qy     601 caaaacaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattat 660
Db     601 CAAAACAAAGTTGTATTGTATGGCCTAAATTACAGCACACTGACACCACACGTATATTAT 660

Qy     661 tctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagta 720
Db     661 TCTCTCTCCATTATCACAGGATGTAACTGTAAAAATTTTGTATGTTAAACATTTGTAGTA 720

Qy     721 aatattgctagcatttacgtctacggaatttattgaaaaaatgtagtattgttttatata 780
Db     721 AATATTGCTAGCATTTACGTCTACGGAATTTATTGAAAAAATGTAGTATTGTTTTATATA 780

Qy     781 attttaataaaactgtaaatcgtctggcttcgtttctggatggaggataaatagtgaata 840
Db     781 ATTTTAATAAAACTGTAAATCGTCTGGCTTCGTTTCTGGATGGAGGATAAATAGTGAATA 840

Qy     841 cgaatgggaaacaccacaacaaccacgccgctgcgttctgcgaatcacatgagcgatcag 900
Db     841 CGAATGGGAAACACCACAACAACCACGCCGCTGCGTTCTGCGAATCACATGAGCGATCAG 900

Qy     901 tgccttgctgttccgtgaacttgcacgcaaggacgagagcctttctgcctttgcatgcaa 960
Db     901 TGCCTTGCTGTTCCGTGAACTTGCACGCAAGGACGAGAGCCTTTCTGCCTTTGCATGCAA 960

Qy     961 ggacaagagtctttacatgcaaggacaaataactcccacgcgccccaccgtgctttggca 1020
Db     961 GGACAAGAGTCTTTACATGCAAGGACAAATAACTCCCACGCGCCCCACCGTGCTTTGGCA 1020

Qy    1021 agccacatggcaccctgccgatcacaattcacaggcccaggcttccggtggtcgcgtgcc 1080
Db    1021 AGCCACATGGCACCCTGCCGATCACAATTCACAGGCCCAGGCTTCCGGTGGTCGCGTGCC 1080

Qy    1081 gtgagtctgacaccgcaccacatggccgccgtaggccgtgcctacgcaccaaggcgactc 1140
Db    1081 GTGAGTCTGACACCGCACCACATGGCCGCCGTAGGCCGTGCCTACGCACCAAGGCGACTC 1140

Qy    1141 gtggtgccaggctctcggcggctttggagtcggtgccatgccgcggggtccgtggaccgc 1200
Db    1141 GTGGTGCCAGGCTCTCGGCGGCTTTGGAGTCGGTGCCATGCCGCGGGGTCCGTGGACCGC 1200

Qy    1201 tcctaggggccaggacgaagctgcaccgcacaagcgggccgcgcgtgactccgtgaccgt 1260
Db    1201 TCCTAGGGGCCAGGACGAAGCTGCACCGCACAAGCGGGCCGCGCGTGACTCCGTGACCGT 1260

Qy    1261 gaggcgggcgtaaccaggagcttccgccacgcttgagaccacgtgacggcgcagaggagc 1320
Db    1261 GAGGCGGGCGTAACCAGGAGCTTCCGCCACGCTTGAGACCACGTGACGGCGCAGAGGAGC 1320

Qy    1321 tccacgcgatcaaaagcgcccgccacttctaaaggtcaggggtcttgcgttctgccctc 1380
Db    1321 TCCACGCGATCAAAAGCGCCCGCCACTTCTAAAGGTCAGGGGTCTTGCGTTCTGCCCCTC 1380

Qy    1381 gtgcttccttcaaattctggacctagtggatcaatttacgtacacctcagcaaccgatgc 1440
Db    1381 GTGCTTCCTTCAAATTCTGGACCTAGTGGATCAATTTACGTACACCTCAGCAACCGATGC 1440

Qy    1441 agccagtatgatgagcacgattgtgacgtgttgggggtcatggtcaatggcaaccgagca 1500
Db    1441 AGCCAGTATGATGAGCACGATTGTGACGTGTTGGGGGTCATGGTCAATGGCAACCGAGCA 1500

Qy    1501 cgaattggtagtgtctgctttttgtacacgtgatagcatttgattcgttcattcaatttg 1560
Db    1501 CGAATTGGTAGTGTCTGCTTTTTGTACACGTGATAGCATTTGATTCGTTCATTCAATTTG 1560

Qy    1561 aactgtttaaacttatatatgtagagaaattagtccaactcatgcttaataaaaagtata 1620
Db    1561 AACTGTTTAAACTTATATATGTAGAGAAATTAGTCCAACTCATGCTTAATAAAAAGTATA 1620

Qy    1621 aaacccatcgaatttatgaattatgatagcaggtatcctatccattgtcatcgctcacag 1680
Db    1621 AAACCCATCGAATTTATGAATTATGATAGCAGGTATCCTATCCATTGTCATCGCTCACAG 1680

Qy    1681 tcacagaggtagccactgccgacggccgacggcctcccatttcgctccccctcctactcct 1740
Db    1681 TCACAGAGGTAGCCACTGCCGACGGCCGACGGCCTCCCATTTCGCTCCCCTCCTACTCCT 1740

Qy    1741 atgctgcggtccagcaaaagttcgggccttccggcaatccgccggcgcccgtcggctcaa 1800
Db    1741 ATGCTGCGGTCCAGCAAAAGTTCGGGCCTTCCGGCAATCCGCCGGCGCCCGTCGGCTCAA 1800

Qy    1801 atcgcatctaccgcggctagaagctctctcttcctccctccgatccggtggggtccattt 1860
Db    1801 ATCGCATCTACCGCGGCTAGAAGCTCTCTCTTCCTCCCTCCGATCCGGTGGGGTCCATTT 1860

Qy    1861 ccttcaattgtggcagtggccgtctcgaaccctctataaatccccccaccccggacaccct 1920
Db    1861 CCTTCAATTGTGGCAGTGGCCGTCTCGAACCCTCTATAAATCCCCCACCCCGGACACCCT 1920

Qy    1921 tccccgaccacacagcccaacaacaaggagcgcggcggcccctccttccttcctcccact 1980
Db    1921 TCCCCGACCACACAGCCCAACAACAAGGAGCGCGGCGGCCCCTCCTTCCTTCCTCCCACT 1980

Qy    1981 tctctcgcgcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagt 2040
Db    1981 TCTCTCGCGCGGCGCTCGCTTACCTCGCCTCGCATTCCGTTCGAGCAGGGGAGCGGCAGT 2040

Qy    2041 gagaagggagggaattaaggcaaccatgg 2069
```

```
          ||||||||||||||||||||||||||||||
Db   2041 GAGAAGGGAGGGAATTAAGGCAACCATGG 2069


RESULT   2
AX151207
LOCUS       AX151207      1959 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 1 from Patent WO0140440.
ACCESSION   AX151207
VERSION     AX151207.1  GI:14533375
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1959)
  AUTHORS   Armstrong,K., Hey,T.D., Folkerts,O., Smith,K.A. and Hopkins,N.L.
  TITLE     Maize mip synthase promoter
  JOURNAL   Patent: WO 0140440-A 1 07-JUN-2001;
            Dow AgroSciences LLC (US)
FEATURES             Location/Qualifiers
     source          1. .1959
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
     CDS             137. .1669
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC42556.1"
                     /db_xref="GI:14533376"
                     /translation="MFIESFRVESPHVRYGPTEIESEYRYDTTELVHEAKDGASRWVV
                     RPKSVKYNFRTSTAVPKLGVMLVGWGGNNGSTLTAGVIANREGISWATKDKVQQANYY
                     GSLTQASTIRVGSYNGEEIYAPFKSLLPMVNPDDLVFGGWDISSMNLADAMTRAKVLD
                     IDLQKQLRPYMESMVPLPGVYDPDFIAANQGSRANNVIKGTKKEQVEQIIKDIREFKE
                     KNKVDKVVVLWTANTERYSNVCAGLNDTMENLLASVDKNEAEISPSTLYAIACVTEGV
                     PFINGSPQNTFVPGLIDLAIKNNCLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIVS
                     YNHLGNNDGMNLSAPQTFRSKEISKSNVVDDMVSSNAILYGPGEHPDHVVVIKYVPYV
                     GDSKRAMDEYTSEIFMGGKSTIVLHNTCEDSLLAAPIILDLVLLAELSTRIQLKPEGT
                     DKFHSFHPVATILSYLTKAPLVPPGTPVVNALAKQRAMLENIMRACVGLAPENNMILE
                     YK"
BASE COUNT      472 a    518 c    568 g    401 t
ORIGIN


  Query Match             5.8%;  Score 119;  DB 6;  Length 1959;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-16;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1945 aaggagcgcggcggcccctccttccttcctcccacttctctcgcgcggcgctcgcttacc 2004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17 AAGGAGCGCGGCGGCCCCTCCTTCCTTCCTCCCACTTCTCTCGCGCGGCGCTCGCTTACC 76

Qy   2005 tcgcctcgcattccgttcgagcaggggagcggcagtgagaagggagggaattaaggcaa 2063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 TCGCCTCGCATTCCGTTCGAGCAGGGGAGCGGCAGTGAGAAGGGAGGGAATTAAGGCAA 135


RESULT   3
AX090098
LOCUS       AX090098      753 bp    DNA             PAT       21-MAR-2001
DEFINITION  Sequence 43 from Patent WO0116307.
ACCESSION   AX090098
VERSION     AX090098.1  GI:13444059
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 753)
  AUTHORS   Conner,T.W. and Tzafrir,I.
  TITLE     Plant regulatory sequences for selective control of gene expression
  JOURNAL   Patent: WO 0116307-A 43 08-MAR-2001;
            Renessen LLC (US)
FEATURES             Location/Qualifiers
     source          1. .753
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
BASE COUNT      188 a    243 c    131 g    191 t
ORIGIN


  Query Match             5.1%;  Score 105.2;  DB 6;  Length 753;
  Best Local Similarity  69.4%;  Pred. No. 3.1e-13;
  Matches  247;  Conservative    0;  Mismatches   68;  Indels   41;  Gaps    6;

Qy   1696 ctgccgacggccgacggcctcccatttcgctcccctcctactcctatgctgcggtccagc 1755
          | |||| | ||||||| ||||||||||||||||||     |||||  ||||||||| |||
Db    410 CGGCCGCCAGCCGACGTCCTCCCATTTCGCTCCCC-----CTCCTCCGCTGCGGTCGAGC 464

Qy   1756 aaaagttcgggccttccggcaatccgccggcgcccgtcggctcaaatcgcat-------- 1807
          ||||||||| |||| |||||||||||| || || |||| ||| ||||| || ||
Db    465 AAAAGTTCCGGCCATCCGGCAATCCCCCCGCACCCGGCGGTTCAAACCGTATCTTTCTGA 524

Qy   1808 -ctaccgcggctagaagctctctcttcc--tccctccgatccggtggggtccatttcctt 1864
           ||  ||||||||  |  || ||| |||  ||||| ||||||||||||||||| |||| ||
Db    525 CCTGACGCGGCTACGACGTCGCTCCTCCGGTCCCTTCGATCCGGTGGGGTCCGTTTCTTT 584

Qy   1865 caattgtggcagtggccgtc-----------------tcgaaccctctataaatccccc 1906
          |||  | |||   |   | |                 ||||||||||||||||||||
Db    585 CAAGCGCGGCCTCGCTGGCCGCCTCGTGGCAGTGACCGTCGAACCCTCTATAAATCCCGT 644

Qy   1907 accccggacaccctccccgaccacacagcccaacaacaaggagcgcggcggcccctcct 1966
          |||||  |||||||||| ||| |||||| ||      ||| |    |   |  ||||||
Db    645 GCCCCGAGCACCCTTCCTCGATCACACAACC------CAAAGCAGCCACAGCAGCCTCCT 698

Qy   1967 tccttcctcccacttctctcgcgcggcgctcgcttacctcgcctcgcattccgttc 2022
          |||| |   |    || ||||||| ||||||| |||||||| |||||||| |||
Db    699 TCCTCCTCTCACTCTCGCTCGCGCTGCGCTCGC-TACCTCGCTTCGCATTCCATTC 753


RESULT   4
AR148903
LOCUS       AR148903      991 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 7 from patent US 6225529.
ACCESSION   AR148903
VERSION     AR148903.1  GI:15112993
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 991)
  AUTHORS   Lappegard,K.K. and Martino-Catt,S.J.
  TITLE     Seed-preferred promoters
  JOURNAL   Patent: US 6225529-A 7 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .991
                     /organism="unknown"
BASE COUNT      266 a    259 c    183 g    283 t
ORIGIN


  Query Match             4.9%;  Score 101.6;  DB 6;  Length 991;
  Best Local Similarity  68.9%;  Pred. No. 2e-12;
  Matches  230;  Conservative    0;  Mismatches   64;  Indels   40;  Gaps    5;

Qy   1696 ctgccgacggccgacggcctcccatttcgctcccctcctactcctatgctgcggtccagc 1755
          | |||| | ||||||| ||||||||||||||||||     |||||  ||||||||| |||
Db    659 CGGCCGCCAGCCGACGTCCTCCCATTTCGCTCCCC-----CTCCTCCGCTGCGGTCGAGC 713

Qy   1756 aaaagttcgggccttccggcaatccgccggcgcccgtcggctcaaatcgcat-------- 1807
          ||||||||| |||| |||||||||||| ||||| |||| ||| ||||||||| ||
Db    714 AAAAGTTCCGGCCATCCGGCAATCCCCCGGCACCCGGCGGTTCAAATCGTATCTTTCTGA 773

Qy   1808 -ctaccgcggctagaagctctctcttcc--tccctccgatccggtggggtccatttcctt 1864
           ||  ||||||||  |  || ||| |||  |||||||||||||||||| |||| ||
```

```
Db     774 CCTGACGCGGCTACGACGTCGCTCCTCCGGTCCCTTCGATCCGGTGGGGTCCGTTTCTTT 833

Qy    1865 caattgtggcagtggccgtc------------------tcgaaccctctataaatccccc 1906
           |||  | |||    |   | |                  ||||||||||||||||||
Db     834 CAAGCGCGGCCTCGCTGGCCGCCTCGTGGCAGTGACCGTCGAACCCTCTATAAATCCCGT 893

Qy    1907 accccggacacccttccccgaccacacagcccaacaacaaggagcgcggcggcccctcct 1966
            |||||  |||||||| ||| |||||| ||          |||||    |   |  ||||||
Db     894 GCCCCGAGCACCCTTCCTCGATCACACAACC------CAAGGCAGCCACAGCAGCCTCCT 947

Qy    1967 tccttcctcccacttctctcgcgcggcgctcgct 2000
            |||| |    |      || |||| || |||||||||
Db     948 TCCTCCTCTCACTCTCGCTCGTGCTGCGCTCGCT 981


RESULT   5
AE001398/c
LOCUS       AE001398     14867 bp    DNA             INV       06-NOV-1998
DEFINITION  Plasmodium falciparum chromosome 2, section 35 of 73 of the
            complete sequence.
ACCESSION   AE001398 AE001362
VERSION     AE001398.1  GI:3845197
KEYWORDS    .
SOURCE      malaria parasite P. falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 14867)
  AUTHORS   Gardner,M.J., Tettelin,H., Carucci,D.J., Cummings,L.M., Aravind,L.,
            Koonin,E.V., Shallom,S., Mason,T., Yu,K., Fujii,C., Pederson,J.,
            Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Pertea,M.,
            Salzberg,S., Zhou,L., Sutton,G.G., Clayton,R., White,O.,
            Smith,H.O., Fraser,C.M., Hoffman,S.L. et,al.
  TITLE     Chromosome 2 sequence of the human malaria parasite Plasmodium
            falciparum
  JOURNAL   Science 282 (5391), 1126-1132 (1998)
  MEDLINE   99021743
  REMARK    Erratum:[[published erratum appears in Science 1998 Dec
            4;282(5395):1827]]
REFERENCE   2  (bases 1 to 14867)
  AUTHORS   Gardner,M.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1998) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20814, USA
FEATURES             Location/Qualifiers
     source          1..14867
                     /organism="Plasmodium falciparum"
                     /db_xref="taxon:5833"
                     /chromosome="2"
     gene            complement(1570..2424)
                     /gene="PFB0490c"
     CDS             complement(1570..2424)
                     /gene="PFB0490c"
                     /note="predicted by GlimmerM"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC71887.1"
                     /db_xref="GI:3845198"
                     /translation="MKEKNEKIMDYLSCPLDDVVDREKKSGKNSLLKSSSTKKSDYKK
                     SSIFSKKRDSHKKGSSFRGRRSGFINRKSGSFKKPYYNNRLINKNYNNYKGRNFHNGR
                     DNFKGRTGSFGSRVFDNRKGSFKKRFISNRNKSSVKSYRGNGSNKMGRKSFNKAPTSR
                     TVVTKRLNNYKTVSAPVKKFNNLNISLYRKNRTFALNTKRSKPVGTIKSSVPRKRIKK
                     GLKKGSLKSKTRKSTSGSKFKPLNKYFLSKIKIVTSLNKIPSPLKEQKNTEVNLPESL
                     NNATTKKN"
     gene            11241..14606
                     /gene="PFB0495w"
     CDS             11241..14606
                     /gene="PFB0495w"
                     /note="predicted by GlimmerM"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC71888.1"
                     /db_xref="GI:3845199"
                     /translation="MFALKKNTVREGFVNICFSYLKKLYLKSNFVTVNLNYETNNEKR
                     TNKKIYKKSKAQSLFDKGLNIHDKLILFKNLPKYKCAKYECISAKEVYKYLLDEYKKC
                     FNYISLCDIIQSVKIFDELDKTFTDYNFYIEVKNIDKNVLNKINEIYFKNKDITFHRR
                     EILGKICNKIMSYIHEMNGNELIHFLIYFFRWNKNDKNLILFYNYYFNYVFDHMYLFN
                     HEIYKLLFIFNKYLNNNSNIPFNKNLIQEMEFNLYYFREIKNEKNYIIKMNKKEIYKK
                     CFAKFHENVDHIDNEKILNILRLYVDNSILDIDINNKMLCNLNNNLINENIEYISKLL
                     NFYCTLIKKGKYDNDMTIYKLKEVIKATHHILCDKTKNLETFCSDIDYSTLLNSLNNK
                     FILNKIIDKNFILFYECLLKILLNIKFVNFQSLCISLISLKNIYYNILRNNVYIVNNV
                     LFNDIMKFSLYLCNIFLGKRIKTENENAVLIIHNNDQTNYSNKENIKDIIIQKRIKEY
                     IFYKMENYKDFHFKLKDSDLLSIKLLSNTFVKINEVYNSYDFYLLFNNISCILYNFLV
                     NRNSVKKYKDTYIYILNDLSFVYKYIKNNDRTKKKKNFFLLSSSMKELICKNILSVSN
                     RYIKHLHEEDNFDQKDQYVCSLTFLNNLFFDKIIHFHYIYNLWCHVYKTYNYFKCNKL
                     INEDIISLLLLTCSKFQYFIENNSNDRYCRKELIHLKYNIIDDLIKNYLNTYKSISID
                     NISKIFISLSNSKYTCEVNENLLLESLQSEFEKVTKTSKKGGIHMMDNNLLDNNNSCE
                     KYEHRYIEYKKENLFINLNKIIECLIKLNIFLYLKKKKTYLYLYKQSLCPINLKENIL
                     KKILYIANNLYMYEMYGYVCEMLERVLSSHKEQNLFSYNYNKNVEHKMFDKILCHISE
                     DDYIEMSNTMYVLFYDYLKNINSERQSNILRNNSTNDRFIDEIKEKKYKLNNNTLIKH
                     NNVKLNYEKSNNSNGNISNILKDDKNKNHNNVEMDLIDNKNENKKIQEKGQNGENCEN
                     CKDVLVNDIINIFGFLKMEKKKFLFFQLYMYLCNITKFKRRYVSSSSLFHMDVFKIIK
                     DMNLKYLCLENYKIKNEECAFLYTIDIVLFKER"
BASE COUNT     6284 a   1019 c   1106 g   6458 t
ORIGIN

  Query Match             3.0%;  Score 62.6;  DB 3;  Length 14867;
  Best Local Similarity   46.1%;  Pred. No. 0.0012;
  Matches  352;  Conservative    0;  Mismatches  404;  Indels    7;  Gaps    4;

Qy      27 agtaaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagc 86
           | |||||||  |||    |||  |   |||  |  | |||| | ||     | | |     ||
Db    8397 ATTAAATATATAATTTTAATTAATTAATTATAAAAAATAAAATAAATAGACTCATAAATA 8338

Qy      87 aactaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaa-acatctgcca 145
            |  || |   || |           | | |    || |    | || ||    ||    ||      |
Db    8337 ATATATAATAATTTAATTAAAAAATAATTTAATATATTTTAATTTAATTAATTAATTA 8278

Qy     146 agaatacaattcttttatacacaacttatatgtgagttcttttctcttgtaactcttat 205
            |  ||| |  |  | ||||| | |   | |||   | | |    |||| | |    ||    | ||
Db    8277 AATATATATTAATATTATATATATATCATAATTTAATATATTTTTTTATAAATAATAAT 8218

Qy     206 taataaaacattttggctattaaataatggcaactaagttagcaccactgtaattagat 265
           |||| ||   |   |||          | |||  |   | ||| |||       | | ||  | ||
Db    8217 TAATTAATTAAGTTT---ATTAAAAATAAATATATTAATTTAATTTAAATATATATAAAT 8161

Qy     266 tttgtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaag 325
              | | |    | || ||    | | || |        | |   |  |  | |  | |||      || ||
Db    8160 AATATTTAATATAAATTAAATTAATATAATTAATTATTTATTATTCATTTAAATAAAAAA 8101

Qy     326 tagaaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaact 385
           || ||||  ||      |         | | |    ||    |    |     |     |   | | || || |
Db    8100 TAATAAATTAAATATATATAATATATATATTTATATGTAAATATATTATAATTATAATT 8041

Qy     386 gaaaacattagcttttttatagctctcgctctctgctagtatgtgttataaaatcattta 445
               |    |||   |  |||        |          |  |    |   |||      ||||   ||      ||   |
Db    8040 ATTATTTTTATTTAATTAATTAATTTATATTATTATTTTATTATTTATTTAACATTTAAA 7981

Qy     446 ccaattaccttttttaaatatactgtacgtagtttcatcagtagaactactcacggagctaa 505
             |    |     || || |||   |         || || ||  |  |     |  ||| |     |   |||
Db    7980 AAATAAATAATTATATATATATAATAATAATTAAATAAAATAATTTAATTATAAAAATAA 7921

Qy     506 aacaaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccg-tgagc 564
           ||  |||||    |    | | ||  | ||| |   ||             || | |  |     |   |   |
Db    7920 AATTAAAAAGAATTATTTATTTATATACACATTTTATATATTTAAATATAATAGTTTTAT 7861

Qy     565 taaagtccaaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggc 624
           ||   | || || ||| |   |         | |||||      -|| |  |  || |||| |||
Db    7860 TATTATACATAACAAATAGATTAATTAAAATAATATATTAATATATTTATATTTTATTAT 7801

Qy     625 ctaaattacagcacactgacaccacacgtatattattctctctccattatcacaggatgt 684
            |   ||||  |       | |     | |   |   | ||||      |          |   | |    |
Db    7800 TTTTATTATAAATTAAT--AATTATATTTTAAATATTAAATTATATAAAATATATATATT 7743

Qy     685 aactgtaaaaatttttgtatgttaaacatttgtagtaaatattgctagcatttacgtctac 744
             |   | ||  | |   |    ||   |   ||     |   | | || ||  ||| |  |   |
```

```
Db   7742 TATAATTAATTTATAATTAATTTTAATTTAAAATATATTTTTATTATTATTAAATTAAAT 7683

Qy    745 ggaatttattgaaaaaatgtagtattgttttatataattttaa 787
          || | | | |||||  || || ||||  ||||| | |  ||||
Db   7682 GGTAATAAATGAAATTATTTATTATTAATTTATTTTAAGTTAA 7640


RESULT   6
I66494
LOCUS       I66494       7218 bp    DNA             PAT       28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .7218
                     /organism="unknown"
BASE COUNT     1944 a   1491 c   1486 g   1929 t    368 others
ORIGIN


  Query Match             3.0%;  Score 62.4;  DB 6;  Length 7218;
  Best Local Similarity   0.8%;  Pred. No. 0.0013;
  Matches    3;  Conservative  234;  Mismatches  135;  Indels    0;  Gaps    0;

Qy   1652 ggtatcctatccattgtcatcgctcacagtcacagaggtagccactgccgacggccgacg 1711
          | | :::: ::: :: :: :: :: ::: :  :: :     :  :: :: ::  :  ::  :
Db   1064 GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123

Qy   1712 gcctcccatttcgctcccctcctactcctatgctgcggtccagcaaaagttcgggccttc 1771
          :::::: :::: :::::::::: ::::: : :: :  :::  :     :::   :::::
Db   1124 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1183

Qy   1772 cggcaatccgccggcgcccgtcggctcaaatcgcatctaccgcggctagaagctctctct 1831
          :  :  ::: ::  : ::: ::  :::   :: : ::: :: : ::   :::::::::
Db   1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243

Qy   1832 tcctccctccgatccggtggggtccatttccttcaattgtggcagtggccgtctcgaacc 1891
          :::::::::: :::  :   ::: :::::::: :: :  :  :  :: :::   ::
Db   1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303

Qy   1892 ctctataaatcccccaccccggacaccctttccccgaccacacagcccaacaacaaggagc 1951
          :::: :   :::::: ::::   : ::::::::: :: : :  :::  :  :     :
Db   1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363

Qy   1952 gcggcggccccctccttcctcccacttctctcgcgcggcgctcgcttacctcgcctc 2011
          :  :  :::::::::::::::::::: :::::::: : :  : ::: ::: :::: ::::
Db   1364 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1423

Qy   2012 gcattccgttcg 2023
            : :::: :::|
Db   1424 YYYYYYYYYYYG 1435


RESULT   7
AX083744
LOCUS       AX083744     1141 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 22 from Patent WO0111061.
ACCESSION   AX083744
VERSION     AX083744.1  GI:13185472
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1141)
  AUTHORS   Kunst,L. and Clemens,S.
  TITLE     Regulation of embryonic transcription in plants
  JOURNAL   Patent: WO 0111061-A 22 15-FEB-2001;
            UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES             Location/Qualifiers
     source          1. .1141
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     promoter        1. .1141
                     /note="consensus sequence of A.t., L.a., and B.n. FAE1
                     promoters"
BASE COUNT     123 a     32 c     42 g    112 t    832 others
ORIGIN


  Query Match             2.9%;  Score 61;  DB 6;  Length 1141;
  Best Local Similarity   12.8%;  Pred. No. 0.0026;
  Matches  109;  Conservative  281;  Mismatches  457;  Indels    6;  Gaps    1;

Qy     32 atatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaacta 91
          :      |: :| :  :: : : :: :    ::  | ::: : :|    ||:|  :: ::
Db    278 MNNNNNNAKAMCKRAKYWGWNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWW 337

Qy     92 aataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaata 151
           : ||:   :      :|:: |:: |::: :::  : :::  ||:   ::    :: |:
Db    338 KWSAAMGVYWNNNNNNNWTYKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTM 397

Qy    152 caattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaataa 211
          :     || :  : |  :  :::::||          |::  |: : : |:::::     :
Db    398 VNNNNGTTMWKRMWAWYWKMDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKW 457

Qy    212 aacatttttggctattaaataatggcaactaagttagcaccactgtaattagattttgtc 271
          :| :::| |   | :: : | :   :: :   |: :::|  ::| :::::::  :
Db    458 RAWDHKTCTHNNTTWWKMKTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRRYAHANNNN 517

Qy    272 tggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtagaaa 331
          :: :: | |:: ::    ::   |  : :      ::: ::|  : ::::: : :|:|
Db    518 WDYWWKACTWYKYBVCSKWWNNYAAWYTKSSWNYTSRYYRWKTNNSWRWRSDTRSMGRAN 577

Qy    332 atggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaaaac 391
          : :| ::|::  | :: : :   ::   | ::  :   : :|: :  :  | : | |
Db    578 NYARABHYGYKWNTRWWBWSHTWBHBRAGAAHYWMBMMYBAKCHCMKAWYKAKKYAGAGG 637

Qy    392 attagctttttatagctctcgctctctgctagtatgtgttataaaatcatttttaccaatt 451
                            ||  : ::  :  :  :   ::: :  ||  |::|:|  ::
Db    638 SNNNNNNNNNNNNNNNATCARDDYYAASRWYAMANAKWYYYKBAANNAYYTHANNWWGC 697

Qy    452 acctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaacaaa 511
          :   |:| : ::          ||::              |  :|       | ::|| ::::
Db    698 WNNATDTRRTMWKNNNNNNAGTWKNNNNNNNAKNASAAKNYA------AAAVKAAKKHWRW 751

Qy    512 aaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaagtc 571
          |  :   |:: :|  : |: :   |  : |       | : ::|    ::|:
Db    752 ANKWAMRGWHADAAABTTDKRNNGAYTKYTTTNNNNTYRGVVTNTAARDGWANNNNNNNN 811

Qy    572 caaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaaatt 631
                : ::::|   :  :            |:|::  : : ||:::::: :|
Db    812 NNNNNNGWSDMWVTWWAYANYGTNNNNNNNNNNAYAWWTNKWYYTTDDRWRBAYTNNNNN 871

Qy    632 acagcacactgacaccacacgtatattattctctctccattatcacaggatgtaactgta 691
          :  :  |: :::|:::::  :|::::: |     | :      :         : |::
Db    872 NRMAYYGAYADDYAYYMSDTCDAWMKWDATKMNNATTYNRGTAWRTNNNNNNMTMKTKYY 931

Qy    692 aaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaattt 751
           :||:       :   |: ::  :  :|   :   |   ::  | ::|::   :||:
Db    932 BHAAWNNNNNNGKMCTAHTWWVCKATKTTKGCWMNCTTTCRKYKNNCTWYTWMTTTRTTW 991

Qy    752 attgaaaaaatgtagtattgttttatataatttttaataaaactgtaaatcgtctggcttc 811
             ::         |:   | ::   ::  ::|::|| :: |::: |: : :  |::
Db    992 YAATRWKTNNATGSMTRCNATGWKNNNYWTGWKTRWTAYRMATRWMKAWWKVMATGSWNT 1051

Qy    812 gtttctggatggaggataaatagtgaatacgaatgggaaacaccacaacaaccacgccgc 871
          :  :   |:  :|::  | |   ::        : |:: ||  ::: :  :   :|:|
Db   1052 NSYARWAYKTRAYKGWYYNACAWRWRWGKATCYMTDNAWWTACATSWMATHKYNWHMCKC 1111
```

```
Qy    872 tgcgttctgcgaa 884
                : :|
Db   1112 NNNNNNNNTMMRA 1124


RESULT   8
AL354718
LOCUS       AL354718   175271 bp    DNA             PRI       31-MAY-2001
DEFINITION  Human DNA sequence from clone RP11-187C18 on chromosome 9, complete
            sequence.
ACCESSION   AL354718 AC027516
VERSION     AL354718.10  GI:13751347
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175271)
  AUTHORS   Sehra,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On May 15, 2001 this sequence version replaced gi:7677898
            gi:13443354.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-187C18 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-187C18.
FEATURES             Location/Qualifiers
     source          1. .175271
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-187C18"
                     /clone_lib="RPCI-11.1"
     repeat_region   443. .874
                     /note="LTR16B repeat: matches 17. .455 of consensus"
     repeat_region   2184. .2357
                     /note="MER45 repeat: matches 4. .178 of consensus"
     repeat_region   4701. .4918
                     /note="L2 repeat: matches 301. .514 of consensus"
     repeat_region   5526. .5759
                     /note="MLT1D repeat: matches 227. .477 of consensus"
     repeat_region   5760. .5972
                     /note="L1MA6 repeat: matches 6086. .6299 of consensus"
     repeat_region   5973. .6220
                     /note="MLT1D repeat: matches 33. .227 of consensus"
     repeat_region   6221. .6573
                     /note="THE1B repeat: matches 1. .364 of consensus"
     repeat_region   6574. .6604
                     /note="MLT1D repeat: matches 1. .33 of consensus"
     repeat_region   6816. .6869
                     /note="27 copies 2 mer tt 75% conserved"
     repeat_region   7939. .8032
                     /note="MIR repeat: matches 103. .197 of consensus"
     repeat_region   8040. .8355
                     /note="AluSx repeat: matches 1. .305 of consensus"
     repeat_region   9192. .9826
                     /note="L1PA8 repeat: matches 5522. .6163 of consensus"
     repeat_region   10472. .10885
                     /note="L1ME1 repeat: matches 4755. .5173 of consensus"
     repeat_region   10894. .12930
                     /note="L1PA7 repeat: matches 4089. .6144 of consensus"
     repeat_region   13015. .13607
                     /note="MER77 repeat: matches 28. .632 of consensus"
     repeat_region   13635. .14266
                     /note="L1ME1 repeat: matches 5263. .5937 of consensus"
     repeat_region   15917. .16183
                     /note="L1M4 repeat: matches 5531. .5801 of consensus"
     repeat_region   20329. .20410
                     /note="41 copies 2 mer at 65% conserved"
     repeat_region   21467. .21881
                     /note="MER74A repeat: matches 119. .558 of consensus"
     repeat_region   22262. .22359
                     /note="L1ME repeat: matches 5544. .5641 of consensus"
     repeat_region   22462. .22760
                     /note="AluSx repeat: matches 1. .299 of consensus"
     repeat_region   23639. .23820
                     /note="MLT1A2 repeat: matches 6. .180 of consensus"
     repeat_region   23829. .23866
                     /note="19 copies 2 mer ct 78% conserved"
     repeat_region   23901. .24046
                     /note="MLT1A2 repeat: matches 238. .388 of consensus"
     repeat_region   24845. .25780
                     /note="L1PA16 repeat: matches 2983. .3906 of consensus"
     repeat_region   25781. .26078
                     /note="AluY repeat: matches 1. .310 of consensus"
     repeat_region   26079. .28287
                     /note="L1PA16 repeat: matches 3906. .6157 of consensus"
     repeat_region   28347. .28864
                     /note="MER74A repeat: matches 1. .532 of consensus"
     repeat_region   29272. .29701
                     /note="MER89 repeat: matches 28. .559 of consensus"
     repeat_region   29702. .29957
                     /note="L1M4 repeat: matches 4006. .4273 of consensus"
     repeat_region   29960. .30009
                     /note="25 copies 2 mer ta 100% conserved"
     repeat_region   30047. .30260
                     /note="MER46A repeat: matches 1. .214 of consensus"
     repeat_region   30380. .30518
                     /note="MER67A repeat: matches 404. .543 of consensus"
     repeat_region   30519. .30798
                     /note="L1PA16 repeat: matches 5869. .6150 of consensus"
     repeat_region   30825. .31252
                     /note="L1M4 repeat: matches 4323. .4755 of consensus"
     repeat_region   31417. .31520
                     /note="AluJo/FRAM repeat: matches 159. .267 of consensus"
     repeat_region   31955. .32776
                     /note="L1ME2 repeat: matches 5345. .6136 of consensus"
     repeat_region   36345. .36566
                     /note="L1ME repeat: matches 5562. .5789 of consensus"
     repeat_region   37479. .37800
                     /note="MER58B repeat: matches 1. .336 of consensus"
     repeat_region   37867. .37917
                     /note="MER58 repeat: matches 2244. .2294 of consensus"
     repeat_region   37967. .38261
                     /note="AluSp repeat: matches 1. .297 of consensus"
     repeat_region   38851. .39159
                     /note="AluSp repeat: matches 1. .313 of consensus"
     repeat_region   41712. .42193
                     /note="L1MA5 repeat: matches 5823. .6300 of consensus"
```

```
     repeat_region   42552. .42775
                     /note="L2 repeat: matches 1589. .1813 of consensus"
     repeat_region   43101. .43156
                     /note="28 copies 2 mer tt 75% conserved"
     repeat_region   43694. .43787
                     /note="L1MC4 repeat: matches 7881. .7977 of consensus"
     repeat_region   43799. .43959
                     /note="L1MC4 repeat: matches 7584. .7728 of consensus"
     repeat_region   43960. .44264
                     /note="AluY repeat: matches 4. .306 of consensus"
     repeat_region   44265. .45123
                     /note="L1MC4 repeat: matches 6608. .7584 of consensus"
     repeat_region   45124. .45419
                     /note="AluY repeat: matches 1. .295 of consensus"
     repeat_region   45425. .45498
                     /note="37 copies 2 mer at 71% conserved"
     repeat_region   45695. .45820
                     /note="L1PA16 repeat: matches 6024. .6157 of consensus"
     repeat_region   48466. .48768
                     /note="AluSq repeat: matches 1. .300 of consensus"
     repeat_region   48940. .49013
                     /note="37 copies 2 mer at 64% conserved"
     repeat_region   49315. .49626
                     /note="156 copies 2 mer aa 54% conserved"
     repeat_region   50885. .51289
                     /note="MER31A repeat: matches 1. .466 of consensus"
     repeat_region   51380. .53498
                     /note="L1PA3 repeat: matches 4028. .6146 of consensus"
     repeat_region   53551. .53921
                     /note="MLT1A1 repeat: matches 1. .365 of consensus"
     repeat_region   53926. .54025
                     /note="L2 repeat: matches 1918. .2019 of consensus"
     repeat_region   54565. .54676
                     /note="L1ME repeat: matches 421. .542 of consensus"
     repeat_region   55322. .55377
                     /note="L1M4 repeat: matches 3856. .3914 of consensus"
     repeat_region   55404. .55604
                     /note="L1M4 repeat: matches 3904. .4086 of consensus"
     repeat_region   55605. .56112
                     /note="MLT1D repeat: matches 1. .503 of consensus"
     repeat_region   56113. .56149
                     /note="L1M4 repeat: matches 4086. .4124 of consensus"
     repeat_region   56467. .56724
                     /note="129 copies 2 mer at 79% conserved"
     repeat_region   56841. .57147
                     /note="AluY repeat: matches 1. .304 of consensus"
     repeat_region   60893. .61141
                     /note="MER8 repeat: matches 1. .239 of consensus"
     repeat_region   61277. .61448
                     /note="MIR repeat: matches 58. .235 of consensus"
     repeat_region   61453. .61874
                     /note="LTR12 repeat: matches 16. .1441 of consensus"
     repeat_region   62227. .62733
                     /note="LTR40a repeat: matches 1. .519 of consensus"
     repeat_region   65994. .66649
                     /note="L2 repeat: matches 2060. .2749 of consensus"
     repeat_region   66776. .66902
                     /note="L2 repeat: matches 1774. .1890 of consensus"
     repeat_region   68482. .68763
                     /note="AluSx repeat: matches 5. .291 of consensus"
     repeat_region   71045. .71331
                     /note="AluSg repeat: matches 1. .288 of consensus"

  Query Match             2.9%;  Score 60;  DB 9;  Length 175271;
  Best Local Similarity   52.4%;  Pred. No. 0.005;
  Matches  132;  Conservative    0;  Mismatches  120;  Indels    0;  Gaps    0;

Qy      9 ttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggagaaattata 68
          |||||| | |      | | | | | ||   | | | | ||| | | | | ||||||
Db  92719 TTTTTTAACTGAGATTCCACTCATAAGGCACCTAAATTTTTAAAGTTAAAACATATTATA 92778

Qy     69 tggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaacctattt 128
          | | |||   ||  ||  |    ||  || | | | | | | ||||||||| ||

Db  92779 TTGACATTAAATTTAAAAAGAATAAACAAGTTAGAGAAGAAATAAAAAAAAAATCTTCAG 92838

Qy    129 ataacaaacatctgccaagaatacaattcttttatacacaacttatatgtgagttctttt 188
          | || ||||  || ||||| |||||| |  |  | |  | ||  |   ||  | || ||
Db  92839 AGAATAAACTGCTTCCAAGTATACAAGTACTAAAGAGGCTACACACTGGTATGCTCATTA 92898

Qy    189 tctcttgtaactcttattaataaaacatttttggctattaaataatggcaactaagttag 248
          |   |   ||| || ||| ||  ||  || |    |    |   ||   || |||  |
Db  92899 TAAATAAAAACACTCATTCATGTAATTTTATCTCATTACTACATATTTAAAATAATCTCT 92958

Qy    249 caccactgtaat 260
           | || | || |
Db  92959 TATCATTATAGT 92970


RESULT   9
AL139261/c
LOCUS       AL139261   128769 bp    DNA             HTG       10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP5-916C24, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AL139261
VERSION     AL139261.8  GI:13274261
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 128769)
  AUTHORS   Plumb,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 11, 2001 this sequence version replaced gi:9931646.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ916C24
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 128305 bases at least Q40
            Consensus quality: 128455 bases at least Q30
            Consensus quality: 128526 bases at least Q20
            Insert size: 128569; sum-of-contigs
            Insert size: 130251; 3.5% error; agarose-fp
            Quality coverage: 8.06x in Q20 bases; sum-of-contigs Quality
            coverage: 8.06x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    33392: contig of 33392 bp in length
            *    33393 33492: gap of      100 bp
            *    33493    40657: contig of 7165 bp in length
            *    40658 40757: gap of      100 bp
            *    40758   128769: contig of 88012 bp in length.
FEATURES             Location/Qualifiers
     source          1. .128769
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP5-916C24"
                     /clone_lib="RPCI-5"
```

```
     misc_feature    1. .33392
                     /note="assembly_fragment:01246
                     fragment_chain:1
                     clone_end:T7
                     vector_side:left"
     misc_feature    33493. .40657
                     /note="assembly_fragment:00130
                     fragment_chain:1"
     misc_feature    40758. .128769
                     /note="assembly_fragment:01367
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    40059 a  24900 c  24955 g  38655 t    200 others
ORIGIN
```

Query Match 2.9%; Score 59.2; DB 2; Length 128769;
Best Local Similarity 38.3%; Pred. No. 0.0074;
Matches 297; Conservative 0; Mismatches 478; Indels 1; Gaps 1;

```
Qy      1 tctagatttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggag 60
Db  40997 TATAAATTTATATAAAAATATATAAAAATATATTATATATAAATATATAAATATATATAA 40938

Qy     61 aaattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaa 120
Db  40937 ATATATTATATATATAATATATATATATATAAATATATTATATATATAATATATAAATAT 40878

Qy    121 acctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtga 180
Db  40877 ATATAAATATATTATATATATAATATATATAATATATTTATATATATTATATATATAATA 40818

Qy    181 gttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaac 240
Db  40817 TATATATTATATTATATATATATAAATATATATAATAAAAATATATTATATATAAATTGA 40758

Qy    241 taagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctat 300
Db  40757 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 40698

Qy    301 ttggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctc 360
Db  40697 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTCAAATATATAATATATTT 40638

Qy    361 acatcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgc 420
Db  40637 ATATATATTATATATATAATATATATATTATATTATATATATATAAATATATAT-AATAA 40579

Qy    421 tagtatgtgttataaaatcattttaccaattacctttttaaataactgtacgtagtttca 480
Db  40578 AAATATATTATATATAAATATATATAAATATATAAATATATATAAATATATATATAAATA 40519

Qy    481 tcagtagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagaga 540
Db  40518 TATATAAATATATAAATATATATAAATATATATATAAATATATATAATATTAAATATATT 40459

Qy    541 tgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacga 600
Db  40458 AAATATAAATATATAAAATATATTATATATAAATATATAAAATATATTATATATAAATAT 40399

Qy    601 caaaacaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattat 660
Db  40398 ATAAAATATATTATATATAAATATATAAAATATATTATATATAAATATATAAAATATATA 40339

Qy    661 tctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagta 720
Db  40338 TGTGTGTGTGTGATTACAGTAAGAAGATGAGTTTTTATATAGGCAGGATGAAAAGCTTTT 40279

Qy    721 aatattgctagcatttacgtctacggaatttattgaaaaaatgtagtattgtttta 776
Db  40278 TTCCTCCTTATGTTTTAGGTTTTTATATTTTTCTTAAACATTTTATTTTTATCCTA 40223
```

```
RESULT  10
AB035880/c
LOCUS       AB035880     1669 bp    DNA             INV       30-MAR-2001
DEFINITION  Chaitoregma tattakana mitochondrial genes for small subunit rRNA,
            tRNA-Val, large subunit rRNA, partial and complete sequences.
ACCESSION   AB035880
VERSION     AB035880.1  GI:13488781
KEYWORDS    tRNA-Val; large subunit rRNA; small subunit rRNA; small subunit
            ribosomal RNA; large subunit ribosomal RNA.
SOURCE      Chaitoregma tattakana mitochondrion DNA.
  ORGANISM  Mitochondrion Chaitoregma tattakana
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
            Aphidiformes; Aphidoidea; Hormaphididae; Chaitoregma.
REFERENCE   1  (sites)
  AUTHORS   Fukatsu,T., Shibao,H., Nikoh,N. and Aoki,S.
  TITLE     Genetically Distinct Populations in an Asian Soldier-Producing
            Aphid, Pseudoregma bambucicola (Homoptera: Aphididae), Identified
            by DNA Fingerprinting and Molecular Phylogenetic Analysis
  JOURNAL   Mol. Phylogenet. Evol. 18 (3), 423-433 (2001)
   PUBMED   11277634
REFERENCE   2  (bases 1 to 1669)
  AUTHORS   Fukatsu,T. and Nikoh,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-1999) Takema Fukatsu, National Institute of
            Bioscience and Human-Technology, Bioengineering Department; 1-1
            Higashi, Tsukuba, Ibaraki 305-8566, Japan
            (E-mail:fukatsu@nibh.go.jp, Tel:81-298-61-6087, Fax:81-298-61-6080)
FEATURES             Location/Qualifiers
     source          1. .1669
                     /organism="Chaitoregma tattakana"
                     /organelle="mitochondrion"
                     /db_xref="taxon:41909"
     rRNA            <1. .426
                     /product="small subunit rRNA"
     tRNA            427. .598
                     /note="putative location"
                     /product="tRNA-Val"
     rRNA            599. .>1669
                     /product="large subunit rRNA"
BASE COUNT      687 a     66 c    146 g    770 t
ORIGIN
```

Query Match 2.7%; Score 55; DB 3; Length 1669;
Best Local Similarity 42.1%; Pred. No. 0.059;
Matches 316; Conservative 0; Mismatches 435; Indels 0; Gaps 0;

```
Qy     49 aaaaatcaggagaaattatatggccatattatagaagcaactaaataaaatgtgcgttgt 108
Db   1186 AAAAAATTATACTAATTTAATCATTATTACTATAAAAAAATTATTTAATTATAAAATTTA 1127

Qy    109 attgaaaaaaaaacctatttataacaaacatctgccaagaatacaattcttttatacaca 168
Db   1126 ATATAAAAAATAAATAAATTATAATAAATTAAATTTAAATAAATAATAAATTTTTACAAA 1067

Qy    169 acttatatgtgagttcttttttctcttgtaactcttattaataaaacatttttggctatta 228
Db   1066 CTTTAAATAAAAATTTCTATTTCTATAAAATATTTAATTTTAAAAATTTATTATATTAAA 1007

Qy    229 aataatggcaactaagttagcaccactgtaattagattttgtctggaacaatttctctga 288
Db   1006 ATTTCAAGCTTATCCCTAAAATAATTTAAATTTAAAATATAATATAAAAAAAAATCAATA 947

Qy    289 ctaagaagctatttggactgtccttttgccaaacaagtagaaaatggaaccgctccttaa 348
Db    946 TAACTTTTAAAATTTTAAATTTAATTTAAATCTTAAATAACTAAATATATTATAACGAAT 887

Qy    349 aaaaccattctcacatcgctgggtgctgaataaaactgaaaacattagctttttatagct 408
Db    886 TAAATTTTTAAACTAAAATTATTTTTTAAATATTTATAAAACAATAAATTAATAATAAAA 827
```

```
Qy    409 ctcgctctctgctagtatgtgttataaaatcattttaccaattacctttttaaataactg 468
          |  |||| |     |        |||||    |  |  ||  |  ||  | ||
Db    826 TTAACTCTATAAATTTTGAGAAATTAAAAAATATAAATTAAAAATTTTAATTAACCCTGA 767

Qy    469 tacgtagtttcatcagtagaactactcacggagctaaaacaaaaaaagttgttctactga 528
          |||  |   |  | | |   | |  |     |  | |   ||      |     |||| |
Db    766 TACAAAAGGTACTAAATTTTAATCTTTTAAAAATTTATTTAATTTCTTTCACAATACTAA 707

Qy    529 taaaagcagagatgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgct 588
          ||||    |    |   || |   |  | |    ||  | ||||||||| |
Db    706 TAAACTATAAATCTAAAAATTAATTTTTACTTAATACTAAAACAAAAAAAAAATAAATTA 647

Qy    589 ccacaataacgacaaaacaaagttgtattgtatggcctaaattacagcacactgacacca 648
          ||||  | |   |||  |  |||  |     | ||||| |  |   | |    |
Db    646 TTTTTATAAATAAATTCAAAACATAAATTTAAAATTATTAATTAAATAAATATAAATAAA 587

Qy    649 cacgtatattattctctctccattatcacaggatgtaactgtaaaaattttgtatgttaa 708
          | ||  |     | ||  |    |    ||       |   |  | | ||| | |
Db    586 GTTATTTAAAACATCTTATCATTGTAAATGATATACTTAAATTTTAGATATTTATTTAAT 527

Qy    709 acatttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaatgtagta 768
          | || | ||  ||||| |   |  ||  |  | |    |  ||||  |||   |    ||
Db    526 AAATATCTATAAAATAATATCAATATAAATTTATTAATATATTATAAAAATCTTTATTTA 467

Qy    769 ttgttttatataatttttaataaaactgtaaa 799
          || ||||  |   | |||||  ||| |  |||
Db    466 TTTTTTTTAACTTTATTAATTAAAATTAAAA 436
```

RESULT 11
AC087568/c
LOCUS AC087568 149733 bp DNA HTG 10-JAN-2001
DEFINITION Pan troglodytes clone RP43-135M11, WORKING DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION AC087568
VERSION AC087568.1 GI:12061434
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 149733)
AUTHORS Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E., Lee-Lin,S.-Q., Legaspi,R., Lim,M., Maduro,Q.L., Maduro,V.B., Masiello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y., Snyder,B., Stantripop,S., Thomas,J.W., Thomas,P.J., Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Zhang,L.-H. and Green,E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 149733)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT ---------- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_mouse@nhgri.nih.gov
---------- Project Information
Center project name: aof
Center clone name: 135M11
---------- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 143625 bases at least Q40
Consensus quality: 144766 bases at least Q30
Consensus quality: 145665 bases at least Q20
Insert size: 150000; agarose-fp
Insert size: 138000; pulse-field-gel
Insert size: 148333; sum-of-contigs
Quality coverage: 9.53x in Q20 bases; agarose-fp
Quality coverage: 10.36x in Q20 bases; pulse-field-gel
Quality coverage: 9.63x in Q20 bases; sum-of-contigs
---------- .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3274: contig of 3274 bp in length
* 3275 3374: gap of unknown length
* 3375 8579: contig of 5205 bp in length
* 8580 8679: gap of unknown length
* 8680 11946: contig of 3267 bp in length
* 11947 12046: gap of unknown length
* 12047 18760: contig of 6714 bp in length
* 18761 18860: gap of unknown length
* 18861 25184: contig of 6324 bp in length
* 25185 25284: gap of unknown length
* 25285 31396: contig of 6112 bp in length
* 31397 31496: gap of unknown length
* 31497 37462: contig of 5966 bp in length
* 37463 37562: gap of unknown length
* 37563 43999: contig of 6437 bp in length
* 44000 44099: gap of unknown length
* 44100 54800: contig of 10701 bp in length
* 54801 54900: gap of unknown length
* 54901 65218: contig of 10318 bp in length
* 65219 65318: gap of unknown length
* 65319 76710: contig of 11392 bp in length
* 76711 76810: gap of unknown length
* 76811 92005: contig of 15195 bp in length
* 92006 92105: gap of unknown length
* 92106 105629: contig of 13524 bp in length
* 105630 105729: gap of unknown length
* 105730 122836: contig of 17107 bp in length
* 122837 122936: gap of unknown length
* 122937 149733: contig of 26797 bp in length.
FEATURES Location/Qualifiers
source 1. .149733
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-135M11"
/clone_lib="RP43"
misc_feature 1. .3274
/note="assembly_fragment"
misc_feature 3375. .8579
/note="assembly_fragment"
misc_feature 8680. .11946
/note="assembly_fragment"
misc_feature 12047. .18760
/note="assembly_fragment"
misc_feature 18861. .25184
/note="assembly_fragment"
misc_feature 25285. .31396
/note="assembly_fragment"
misc_feature 31497. .37462
/note="assembly_fragment"
misc_feature 37563. .43999
/note="assembly_fragment
clone_end:T7
vector_side:left"
misc_feature 44100. .54800
/note="assembly_fragment"
misc_feature 54901. .65218
/note="assembly_fragment"
misc_feature 65319. .76710
/note="assembly_fragment"

```
     misc_feature    76811. .92005
                     /note="assembly_fragment"
     misc_feature    92106. .105629
                     /note="assembly_fragment"
     misc_feature    105730. .122836
                     /note="assembly_fragment"
     misc_feature    122937. .149733
                     /note="assembly_fragment"
BASE COUNT    37083 a  36626 c  36064 g  38478 t   1482 others
ORIGIN


  Query Match             2.6%;  Score 54.2;  DB 2;  Length 149733;
  Best Local Similarity   45.5%;  Pred. No. 0.099;
  Matches  356;  Conservative    0;  Mismatches  413;  Indels   14;  Gaps    4;

Qy      31 aatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaact 90
           |||||    |    || || | ||    |   |||||||    |||||||| |      |
Db   15191 AATATATTTTATATATTTATATATTTATATTTATTATATATTTATATTATATATATTTAT 15132

Qy      91 aaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaat 150
           |    |||| |    |    | ||| | |||   |||||||| | | || |     | |||
Db   15131 ATTATAAATATATTTATATTATAAATATAAAAACATTTATAATATAAATATATTTATAAT 15072

Qy     151 acaattcttttatacacaacttatatgtgagttctttttctcttgta-------actctt 203
           | || | |   |||||    | ||||     |  || |  | |             ||  ||
Db   15071 ATAAATATAAAATACATTTATAATATAAATATATTTATATTATAAATATAAAATACATTT 15012

Qy     204 attaataaaacatttttggctattaaataatggcaactaagttagcaccactgtaattag 263
           || |   ||| || |||      ||||||      ||     ||  |  | | || |||
Db   15011 ATAATATAAATATATTTATACTATAAATATAAAATACATTTATAATATAAATATATTTAT 14952

Qy     264 attttgtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaaca 323
           ||| |   | || |     |   | ||  ||  ||      | |   |||   ||
Db   14951 ATTATAAATATAAAATACATTTATATTATAAATATAAAACACATTTATATTTATTAATAT 14892

Qy     324 agtagaaaatgga---accgctccttaaaaaaccattctcacatcgctgggtgctgaata 380
           |  | |||||  |   |    |  | | | ||  ||    | ||   |   |  | |
Db   14891 AAAATAAAATATATTTATATTTATTAATATAAATATAAATATATTTATATTTATTAATAT 14832

Qy     381 aaactgaaaacattagctttttatagctctcgctctctgctagtatgtgttataaaatca 440
           |||   |||    |   | |||||  | |   | |       | | | |  |   || |
Db   14831 AAATATAAATATATTTATATTTATTAATATAAATATAAATATATTTATATAGTTTTATTA 14772

Qy     441 ttttaccaattacctttttaaataactgtacgtagtttcatcagtagaactactcacgga 500
           | ||   || |   | || |   | | |   || | | |  |  | |  || | |
Db   14771 ATATAAATATAAATATATTTATATAGTTTTATTAATATAAATATAAAATATATTTATATT 14712

Qy     501 gctaaaacaaaa---aaagttgttctactgataaaagcagagatgatgtatgaccgtgac 557
           | ||   |||   ||  |   |  | | ||| ||  | | |  || | |
Db   14711 TATTAATATAAATATAATTTATATAGATTAATATAAATATAAAATATATTTATATAGTTT 14652

Qy     558 cgtgagctaaagtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtatt 617
           | |    || |  |||| | |    |        || |  | |||   ||  | ||||
Db   14651 TATTAATATAAATATAAAATATATTTATATAGTTTTATTAATATAAATATAAAATATATT 14592

Qy     618 gtatggcctaaattacagcacactgacaccacacgtatattattctctctccattatcac 677
             |  ||||   | |   | |  | |  |||||  || | |    || |  |
Db   14591 TATATAGTTTTATTAATATAAATATAAAATATATTTATATAGTTTTATTAATATAAATAT 14532

Qy     678 aggatgtaactgtaaaaattttgtatgttaaacatttgtagtaaatattgctagcattta 737
           |  || ||  | || |  ||| ||   |||| ||    |   | |  |       | ||
Db   14531 AAAATATATTTATATAGTTTTATTAATATAAATATAAAAATATATTTATATGGTTTTATA 14472

Qy     738 cgtctacgga-atttattgaaaaaatgtagtattgttttatataattttaataaaactgt 796
           | ||   | || | || | |||||||| |||  || |||||  ||    |||||||| |
Db   14471 AATATAAAAATATATTTTTATAAAATGTAATATATTTATATATTATAAATATAAAACTAT 14412

Qy     797 aaa 799
           | |
Db   14411 ATA 14409
```

```
RESULT  12
F16N3/c
LOCUS       F16N3      171769 bp    DNA             PLN       31-JUL-1999
DEFINITION  Sequence of BAC F16N3 from Arabidopsis thaliana chromosome 1,
            complete sequence.
ACCESSION   AC007519
VERSION     AC007519.2  GI:5430744
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 171769)
  AUTHORS   Vysotskaia,V.S., Schwartz,J.R., Yu,G., Toriumi,M., Lenz,C., Liu,S.,
            Lee,J., Liu,A., Li,J., Kremenetskaia,I., Gonzalez,A., Altafi,H.,
            Araujo,R., Brooks,S., Buehler,E., Chao,Q., Conn,L., Conway,A.B.,
            Dunn,P., Hansen,N., Huizar,L., Khan,S., Kim,C., Palm,C., Rowley,D.,
            Shinn,P., Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and
            Theologis,A.
  TITLE     Arabidopsis thaliana chromosome 1 BAC F16N3 sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171769)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1999) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
REFERENCE   3  (bases 1 to 171769)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1999) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
REFERENCE   4  (bases 1 to 171769)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1999) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     On Jul 9, 1999 this sequence version replaced gi:4761801.
            The sequence of BAC F16N3 from Arabidopsis thaliana chromosome 1.
FEATURES             Location/Qualifiers
     source          1. .171769
                     /organism="Arabidopsis thaliana"
                     /cultivar="columbia"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="F16N3"
     gene            1144. .2319
                     /gene="F16N3.1"
     CDS             1144. .2319
                     /gene="F16N3.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F16N3.1"
                     /protein_id="AAD46016.1"
                     /db_xref="GI:5668790"
                     /translation="MEQREEKTENIQRKRSRGKSSSSSLPLDLTSEIFSRLPAKSVVR
                     FRCVSKLWSSITTAPYFTNSFETRPNLMFFFKEGDKFFVFTIPQHNQNPNESYSYSSS
                     EIIDSYHTTYPKRCCVTTLTESVHGLICFRKAATPIIWNPTMRKFKPLRKLDERWKNI
                     KVSLGYDPVDGKHKVVCMPYGNAFYECRVLTLGSDQEWRTVKTNHKNSPFTFHGGVCY
                     RQSRCINGVIYYRADTNSGRVILSFDVRSERFDVIELPWDENFGLVMMMSYKGRLACL
                     GFNHEKNSRSMWVLENVEQREWSCHTYLPISHYEPGLENYFNLTGITNDGELIYVPNT
                     VLERFHVIYFDAIRETFRRVIYKGVADKGFRLRNGLEDKPIRRLHFFPNHIETLMSL"
     gene            3391. .4736
                     /gene="F16N3.2"
     CDS             join(3391. .3680,3853. .3997,4153. .4240,4474. .4736)
                     /gene="F16N3.2"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F16N3.2"
                     /protein_id="AAD46017.1"
                     /db_xref="GI:5668791"
                     /translation="MNTFFKLGSLIQRTASQISSSFPKSRFFSDGESAVYHHARLFKK
                     PLSTKLKFNLVNSVSLMGFVDRSIQVMNTGPDRFGVFTILRVKDPLNPNRSFRISLRM
```

```
                WDAMARTCIAHLKLNDHILVSGRLESYSKSSSDVYSGLNLDYQVKVAEVNYVAAPPSH
                VLDSQISKNPKTKTEDDIEESKKDEIYLWQVFFSNPYDWWDNRRNKKNPKQPDFKHKD
                TGEALWLCSDLPDWITRRLELFDQKNRFYDEEKTRRDRLSDYI"
gene            complement(6189. .7979)
                /gene="F16N3.3"
CDS             complement(join(6189. .6974,7590. .7979))
                /gene="F16N3.3"
                /note="Strong similarity to gb|Z15116 serpin (pazx) from
                Hordeum vulgare and is a member of the PF|00079 Serpin
                family.  ESTs gb|R65473, gb|N38150 and gb|AA712968 come
                from this gene."
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46018.1"
                /db_xref="GI:5668792"
                /translation="MDVRESISLQNQVSMNLAKHVITTVSQNSNVIFSPASINVVLSI
                IAAGSAGATKDQILSFLKFSSTDQLNSFSSEIVSAVLADGSANGGPKLSVANGAWIDK
                SLSFKPSFKQLLEDSYKAASNQADFQSKAVEVIAEVNSWAEKETNGLITEVLPEGSAD
                SMTKLIFANALYFKGTWNEKFDESLTQEGEFHLLDGNKVTAPFMTSKKKQYVSAYDGF
                KVLGLPYLQGQDKRQFSMYFYLPDANNGLSDLLDKIVSTPGFLDNHIPRRQVKVREFK
                IPKFKFSFGFDASNVLKGLGLTSPFSGEEGLTEMVESPEMGKNLCVSNIFHKACIEVN
                EEGTEAAAAASAGVIKLRGLLMEEDEIDFVADHPFLLVVTENITGVVLFIGQVVDPLH"
misc_feature    complement(18129. .19068)
                /note="Contains similarity to reverse transcriptase."
misc_feature    complement(21352. .22314)
                /note="Contains similarity to reverse transcriptase."
misc_feature    complement(24661. .25475)
                /note="Contains similarity to polyprotein."
gene            26764. .29416
                /gene="F16N3.4"
CDS             join(26764. .27140,27836. .27969,28100. .28527,28610. .28741,
                28928. .29416)
                /gene="F16N3.4"
                /note="Similar to gb|U39782 lysine and histidine specific
                transporter from Arabidopsis thaliana.  EST gb|Z17527
                comes from this gene."
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46019.1"
                /db_xref="GI:5668793"
                /translation="MDERPETELISIPATPRVSTPEILTPSGQRSPRPATKPSSATWT
                PTSFISPRFLSPIGTPMKRVLVNMKGYLEEVGHLTKLNPQDAWLPITESRNGNAHYAA
                FHNLNAGVGFQALVLPVAFAFLGWSWGILSLTIAYCWQLYTLWILVQLHEAVPGKRYN
                RYVELAQAAFGERLGVWLALFPTVYLSAGTATALILIGGETMKLFFQIVCGPLCTSNP
                LTTVEWYLVFTSLCIVLSQLPNLNSIAGLSLIGAVTAITYSTMVWVLSVSQPRPATIS
                YEPLSMPSTSGSLFAVLNALGIIAFAFRGHNLVLEIQSTMPSTFKHPAHVPMWRGAKI
                SYFLIALCIFPISIGGFWAYGNLMPSGGMLAALYAFHIHDIPRGLLATAFLLVVFSCL
                SSFQIYSMPAFDSFEAGYTSRTNKPCSIWVRSGFRVFFGFVSFFIGVALPFLSSLAGL
                LGGLTLPVTFAYPCFMWVLIKKPAKYSFNWYFHWGLGWLGVAFSLAFSIGGIWSMVTN
                GLKLKFFKPPN"
misc_feature    30699. .31354
                /note="Contains similarity to reverse transcriptase."
gene            complement(40184. .40813)
                /gene="F16N3.5"
CDS             complement(40184. .40813)
                /gene="F16N3.5"
                /note="Similar to gb|U82230 prolamin box binding factor
                (PBF) from Zea mays."
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46020.1"
                /db_xref="GI:5668794"
                /translation="MPSEPNQTRPTRVQPSTAAYPPPNLAEPLPCPRCNSTTTKFCYY
                NNYNLAQPRYYCKSCRRYWTQGGTLRDVPVGGGTRRSSSKRHRSFSTTATSSSSSSSV
                ITTTTQEPATTEASQTKVTNLISGHGSFASLLGLGSGNGGLDYGFGYGYGLEEMSIGY
                LGDSSVGEIPVVDGCGGDTWQIGEIEGKSGGDSLIWPGLEISMQTNDVK"
misc_feature    complement(41293. .45390)
                /note="Strong similarity to gb|X13291 retrotransposon
                Ta1-3 DNA from Arabidopsis thaliana."
gene            46252. .47841
                /gene="F16N3.6"
CDS             join(46252. .46328,46560. .46629,46731. .46790,47011. .47125,
                47393. .47508,47593. .47841)
                /gene="F16N3.6"
                /note="EST gb|AA605495 comes from this gene."
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46021.1"
                /db_xref="GI:5668795"
                /translation="MTLFHFFNCAILTFGPHAVYYSATPLSEYDTLGTSVKAAVVYLA
                TALVKLVCLATFLQVSETEVFDPYQEALKAMIGFIDVAGLYYALAQLTHRNISQNHKF
                QAVGLGWAFADSVLHRLAPLWVGARGLEFTWDYVLQGLEANANLVFTISLAALGSLMW
                LRKNKPKSLIPIIYTCAVIIATMPSITSYLKRVNGWHFPKIVGFELATSLVMAFISCQ
                LFILCQRPSL"
gene            48400. .49970
                /gene="F16N3.7"
CDS             join(48400. .48906,48908. .49708,49740. .49970)
                /gene="F16N3.7"
                /note="Strong simlarity to gb|286426 F10M6.190 cytochrome
                p450 homolog from Arabidopsis thaliana BAC gb|AL021811.
                (May be a pseudogene.)"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46022.1"
                /db_xref="GI:5668796"
                /translation="MASIGLYEAFIGFLCFLISFYFLVKKPFRYLLIKKTHQSYPWNW
                PVLGMLPGILLRLQRIYDYSVEFLENSNMTFQFKGPWFVGMDILATADPTNIHYIMSS
                NFSNYIKGPIFHEIFEAFGDGIINTDAELWRDWRNASQLIFNHQRYQNFSASTTKSKV
                NDGLVPLFNVDLQDVFQRFMFDITFIFLTGTDPRSLSIEMPEIEFAKALDDVGDVIVY
                RHITPRFVWKLQKWIGIGTEKKMSKANATLDRVCEKLIAAKREELRSQGIIDNANGDS
                EDLLTSHIMLDATKYELLNPNDDKFLRDFTVSFMAAGRDSTSSALTWFFWNLTENPNV
                LSKILQEINTNLPRTGSDQDMSSYLNKLVYLHGALNESMRLYPPIPFERKSPIKDDVL
                PSGHKVKSNLNIMIFIYAMGRMKTIWGEDAMEFKPERWISETGGLRHEPSYKFFSFNA
                GPRTCLGKNLAMNLMKTVIVEILQNYEIKIVSGQKIEAKPGLVLHMKHGLKVTMTMKC
                SSLE"
gene            56572. .58134
                /gene="F16N3.8"
CDS             56572. .58134
                /gene="F16N3.8"
                /note="Strong simlarity to gb|286426 F10M6.190 cytochrome
                p450 homolog from Arabidopsis thaliana BAC gb|AL021811."
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAD46023.1"
                /db_xref="GI:5668797"
                /translation="MASIGLYEAFIGFLCFLISFYFLVKKPFSYLLIKKTLQSYPWNW
                PVLGMLPGVLLRLQRIYDCSVEVLENSNMTFQFKGPWFVGMDVLATVDPANIHHIMSS
                NFSNYIKGPIFHEIFEAFGDGIINTDAELWRDWRNASQLIFNHQRYQNFSASTTKTKV
                NDGLVPLFNHFANEEIVVDLEDVFQRFMYDITFIFITGTDPRSLSIEMPEVEFSKALD
                DVGDAIVHRHITPRFVWKLQKWIGIGTEKKMLKAHATFDRVCEKIIAAKREELGSQGI
                TYNSNGEREDLLTSFIKLDATKYEVLKPSHDKFLRDFTIGFMAAGRDSTASTLTWFFW
                NLSKNPNVLTKILQEINTNLPRTGSDQDMSSYLNKLVYLHGALSESMRLYPPIPFQRK
                SPIKEDVLPSGHKVKSNINIMIFIYAMGRMKTIWGEDAMEFKPERWISETGGVRHEPS
                YKFLSFNAGPRTCLGKNLAMNLMKTVIVEILQNYEIKIVSGQKIEPKPGLILHMKHGL
                KVTMTKKCSSLE"
gene            complement(60359. .61414)
                /gene="F16N3.9"
CDS             complement(60359. .61414)
                /gene="F16N3.9"

  Query Match             2.6%;  Score 53.4;  DB 8;  Length 171769;
  Best Local Similarity   49.5%;  Pred. No. 0.15;
  Matches  138;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     127 ttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtgagttctt 186
           || ||  |   ||  |  |||  | ||     | |  |   ||||| | | | |  |||
Db   41194 TTTTACAATTAATGAGAAAAGGCTTCATAATATCTTCAACAAACTTTTCTCTCATCTCTA 41135

Qy     187 tttctcttgtaactcttattaataaaacatttttggctattaaataatggcaactaagtt 246
           || | ||  ||| || | ||||  ||  |  |     | ||||||     |   ||  ||
Db   41134 TTCCGCTAATAATTCATTTTAACGAAGTAAATACTATTTTTAAATCCACTCTCATATTTT 41075

Qy     247 agcaccactgtaattagattttgtctggaacaatttctctgactaagaagctatttggac 306
           | |     | || ||| |||| | | ||| || || | |    || || |||  |
Db   41074 AACCATTTTTTAGTTAAATTTAATATAGAAGAAATTATTTTTAAAAAAAACTAAATATGA 41015
```

```
Qy    307 tgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctcacatcg 366
          |   ||      | |||  ||||||  |     | | ||   | | |||   | ||
Db  41014 ATTAAATTATTTTTAGAAGGGGAAAATAAATTAAATTCATATGTATCAATTAATAAATAA 40955

Qy    367 ctgggtgctgaataaaactgaaaacattagctttttata 405
            ||   ||||| |||||    || |    || | || |
Db  40954 AAGGCAACTGAAAAAAACGTCTAAAAGCCTCTATATAAA 40916


RESULT 13
AX083744/c
LOCUS       AX083744     1141 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 22 from Patent WO0111061.
ACCESSION   AX083744
VERSION     AX083744.1  GI:13185472
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1141)
  AUTHORS   Kunst,L. and Clemens,S.
  TITLE     Regulation of embryonic transcription in plants
  JOURNAL   Patent: WO 0111061-A 22 15-FEB-2001;
            UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES             Location/Qualifiers
     source          1. .1141
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     promoter        1. .1141
                     /note="consensus sequence of A.t., L.a., and B.n. FAE1
                     promoters"
BASE COUNT      123 a     32 c     42 g    112 t    832 others
ORIGIN


  Query Match             2.6%;  Score 53;  DB 6;  Length 1141;
  Best Local Similarity   10.4%;  Pred. No. 0.16;
  Matches   89;  Conservative  326;  Mismatches  438;  Indels    4;  Gaps    2;

Qy      2 ctagatttttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggaga 61
          |: :| |  :  |::::: :  :::::|:::|:|| : ::       : : ::::  :::
Db    912 CYNRAATNNKMATHWMKWTHGAHSKRRTRHHTRTCRRTKYNNNNNNARTVYWYHHAARRW 853

Qy     62 aattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaa 121
          :  ::|:|             :  : :: :: : :                |  :   | |
Db    852 MNAWWTRTNNNNNNNNNNACRNTRTWWABWKHSWCNNNNNNNNNNNNNNNTWCHYTTANAB 793

Qy    122 cctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtgag 181
          :|::     || |:: :||  :::| : | : :: ::|:  |::: :: |  | |   :
Db    792 BCYRANNNNAA-ARMARTCNNYMHAAVTTTHTDWCYKTWMNTWYWDMMTTMBTTTTRNMT 734

Qy    182 ttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaact 241
          |    |                  : |: |:|   :  :   |:  :       : :::|
Db    733 TSTNMTNNNNNNMWACTNNNNNNMWKAYYAHATNNWGCWWNNTDARRTNNTTVMRRRWMT 674

Qy    242 aagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctatt 301
          :| : :   ::: | ||              :  | |  :| : |: |:    :::
Db    673 NTKTRWYSTTRRHHYTGATNNNNNNNNNNNNNNNNSCCTCTRMMTMRWTMKGDGMTVRKK 614

Qy    302 tggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctca 361
          : : |   : : :  :::: | ::: :::|  :    |::: :  : :   | ::
Db    613 VKWRDTTCTYVDVWADSWVWWYANWMRCRDVTYTRNNTYCKSYAHSYWYWSNNAMWYRRY 554

Qy    362 catcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgct 421
          :|   :   ::    :::   :  ::|  ::::      |:|: :: : :   : | ::
Db    553 SARNWSSMARWTTRNNWWMSGBVRMRWAGTMWWRHWNNNNTDTRYYWWWKRWARBTTTVY 494

Qy    422 agtatgtgttataaaatcattttaccaattacctttttaaataactgtacgtagtttcat 481
          ::   :: ::   :: :  :: | :|  : ::|::    |::: : :      :  :
Db    493 DSMCNAKSMWRGNNWRAMKMWWAANNDAGAMDHWTYWMGNNTMWMRRAWKMMNMAWCRRA 434

Qy    482 cagtagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagagat 541
          :          ::: :: :: |:  :  ||     ::  :: : : :: ::    :: ::
Db    433 YCCNNNNNRACVWHKHKMWRWTWKYMWKAACNNNNBKAMYMRVAWMMYSRDTTNTDWMMW 374

Qy    542 gatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacgac 601
            : :::   ::: ::  :        :::|    :::::::: | ||:    :  |:::
Db    373 TSDWBWHWYTVDYTMMRAWNNNNNNNNWRBCKTTSWMWWMDHMNTHCTYGNNTWGSAYBMA 314

Qy    602 aaaacaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattatt 661
          |: : ||    :  :   :   ::   :|        ::::   : :|  :: ::| ::|
Db    313 AMSMWAAGASNBVTYNWCWRMTYMGKTMTNNNNNNKAWYYRTKTVAWCNNRYYYDTAVWT 254

Qy    662 ctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagtaa 721
          ::  :::|  : :: ::  : ::  :: | :: :  :   : : ::::::: : : :
Db    253 BKRNYKYCYAYBWYYBMYMGKHHWBWWRRABHRSWNMWWVKCRNKYMVSWHYHAMRYBKW 194

Qy    722 ata---ttgctagcatttacgtctacggaatttattgaaaaaatgtagtattgttttata 778
          |:|       ::::|::::| |   :  ::: :::: :    |  |  :  ::::::|::
Db    193 ABAVGCNNNWKDRMAHHHWCATNNNMMWWWYAYMHHMHKKGKAAWTNNKTABRDDHBAHV 134

Qy    779 taattttaataaaactgtaaatcgtctggcttcgtttctggatggaggataaatagtgaa 838
          :  :::  ::: |:|:  | |:::|  : :::: : ::    ||::  : ::  : :::
Db    133 KTYWYWRYDYWCAMCWMNAKAKVRTAMKHMWYYTDRYVSANNTGVRWMMMRWCMWWYSMN 74

Qy    839 tacgaatgggaaacacc 855
           :   : |::  : :::
Db     73 NRWYYRMGRKYTWAWMY 57


RESULT 14
AC005505
LOCUS       AC005505   192929 bp    DNA             HTG       12-AUG-2000
DEFINITION  Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
            PROGRESS ***, 1 ordered pieces.
ACCESSION   AC005505
VERSION     AC005505.8  GI:9797710
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      malaria parasite P. falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 192929)
  AUTHORS   Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
            Kurdi,O.B., Conway,A.B. and Davis,R.W.
  TITLE     Plasmodium falciparum 3D7 chromosome 12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 192929)
  AUTHORS   Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
            Center, Stanford University, 855 California Avenue, Palo Alto, CA
            94304, USA
COMMENT     On Aug 12, 2000 this sequence version replaced gi:8810446.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *         1   192929: contig of 192929 bp in length.
FEATURES             Location/Qualifiers
     source          1. .192929
                     /organism="Plasmodium falciparum"
                     /db_xref="taxon:5833"
                     /chromosome="12"
                     /clone="PFYAC181"
                     /clone="3D7"
BASE COUNT    81142 a  17397 c  22111 g  72279 t
ORIGIN


  Query Match             2.5%;  Score 51.6;  DB 2;  Length 192929;
  Best Local Similarity   51.3%;  Pred. No. 0.38;
```

```
  Matches  120;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy      45 atctaaaaatcaggagaaattatatggccatattatagaagcaactaaataaaatgtgcg 104
           || || | ||  | | | | |||||    |||| |||  |   | || ||| |||
Db   96335 ATATATATATGTGTATATAATATATAAAAATATAATACTATATAATATATATAATAATTT 96394

Qy     105 ttgtattgaaaaaaaaacctatttataacaaacatctgccaagaatacaattcttttata 164
           | | ||| |||||||||   |   | || ||| | ||     ||| | || |   || |
Db   96395 TAGGATTTAAAAAAAAAAAAAAAAAAAAAAAAAACTTTGGGAAAAAAAGTAAATTGTT 96454

Qy     165 cacaacttatatgtgagttctttttctcttgtaactcttattaataaaacatttttggct 224
           | |  || | | |   || ||||| | || |   | | ||  |  |    |  ||   |
Db   96455 AAAATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAATATAACATTTTTGATTTTGT 96514

Qy     225 attaaataatggcaactaagttagcaccactgtaattagattttgtctggaaca 278
           ||  | |      |  |   | ||   | || |  ||   ||||||  |  || |
Db   96515 ATCTATTTTAATTATTTTTATAAGATACTCTTTTTTTCTTTTTTGTTTTAAATA 96568


RESULT  15
AC005504
LOCUS       AC005504   104992 bp    DNA             HTG       01-APR-1999
DEFINITION  Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC005504
VERSION     AC005504.3  GI:4558584
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      malaria parasite P. falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 104992)
  AUTHORS   Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B., Conway,A.B.
            and Davis,R.W.
  TITLE     Plasmodium falciparum 3D7 chromosome 12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 104992)
  AUTHORS   Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
            Center, Stanford University, 855 California Avenue, Palo Alto, CA
            94304, USA
COMMENT     On Apr 2, 1999 this sequence version replaced gi:4337172.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    58642: contig of 58642 bp in length
            *    58643    58842: gap of unknown length
            *    58843    91011: contig of 32169 bp in length
            *    91012    91211: gap of unknown length
            *    91212   104992: contig of 13781 bp in length.
FEATURES             Location/Qualifiers
     source          1. .104992
                     /organism="Plasmodium falciparum"
                     /db_xref="taxon:5833"
                     /chromosome="12"
BASE COUNT    44286 a   9326 c   9564 g  41411 t    405 others
ORIGIN


  Query Match             2.5%;  Score 51.2;  DB 2;  Length 104992;
  Best Local Similarity   48.9%;  Pred. No. 0.46;
  Matches  137;  Conservative    0;  Mismatches  143;  Indels    0;  Gaps    0;

Qy     106 tgtattgaaaaaaaaacctatttataacaaacatctgccaagaatacaattcttttatac 165
           | |||  || |||   |||||||| | | |||  | | |  |   || |||| | || |
Db   77042 TATATATAATAAAGTTCCTATTTTTCAAAAAGGTTTCCACAATGTAAAATTTTATTGTGT 77101

Qy     166 acaacttatatgtgagttctttttctcttgtaactcttattaataaaacatttttggcta 225
           ||  |||||| | | |  | |   || | ||     |     | |    | ||  | |
Db   77102 TCATTTTATATTTTAATATTATAGATCCTATATAATATTCATTTCATTGTTCTTAAGGTT 77161

Qy     226 ttaaataatggcaactaagttagcaccactgtaattagattttgtctggaacaatttctc 285
           | ||  |      |  || ||   ||   | ||  || ||||| | |     | || | |
Db   77162 TAAATAATGTATGATAAAATTTATACACATATATATATATTTTTTATTTTTTATTTGCAC 77221

Qy     286 tgactaagaagctatttggactgtccttttgccaaacaagtagaaaatggaaccgctcct 345
           | || | |   |  |  | | |  | ||   | | || || | | |
Db   77222 ACAATATGTAAAAAAATATATTTTTATATTCTTATATAACTACACATTTTTTTTATAAAC 77281

Qy     346 taaaaaaccattctcacatcgctgggtgctgaataaaact 385
           ||||||||||||||||     ||   |    | ||||| |
Db   77282 TAAAAAACCATTCTCATTATTCTAAATAAAAATTAAAATT 77321


Search completed: December 19, 2001, 22:54:10
Job time: 10491 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 21:23:50 ; Search time 199.95 Seconds
(without alignments)
8871.247 Million cell updates/sec

Title: US-09-727-628-3
Perfect score: 2069
Sequence: 1 tctagatttttttttcaattc.....gggaattaaggcaaccatgg 2069

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2069	100.0	2069	22	AAC85921	MIP synthase promo
2	183.6	8.9	936	22	AAF58252	Oligonucleotide D1
3	183.6	8.9	936	22	AAF58254	Oligonucleotide D1
4	183.6	8.9	936	22	AAF58257	Oligonucleotide D1
5	183.6	8.9	936	22	AAF58259	Oligonucleotide D2
6	183.6	8.9	936	22	AAF58262	Oligonucleotide D2
7	183.6	8.9	938	22	AAF58255	Oligonucleotide D1
c 8	182.8	8.8	936	22	AAF58252	Oligonucleotide D1
c 9	182.8	8.8	936	22	AAF58254	Oligonucleotide D1
c 10	182.8	8.8	936	22	AAF58257	Oligonucleotide D1
c 11	182.8	8.8	936	22	AAF58259	Oligonucleotide D2
c 12	182.8	8.8	936	22	AAF58262	Oligonucleotide D2
c 13	182.8	8.8	938	22	AAF58255	Oligonucleotide D1
14	122.2	5.9	483	21	AAC35815	Zea mays DNA fragm
15	119	5.8	1959	22	AAC85922	MIP synthase cDNA.
16	105.2	5.1	753	22	AAD03212	700265029 clone Ze
17	101.6	4.9	991	21	AAZ51583	Maize seed-preferr
c 18	66.4	3.2	244	22	AAF58238	Oligonucleotide D1
19	64.4	3.1	244	22	AAF58238	Oligonucleotide D1
c 20	47.8	2.3	20598	22	AAF24851	Nucleotide sequenc
c 21	47.4	2.3	1069	22	AAH26233	Human prothrombina
c 22	47.4	2.3	1648	22	AAS22816	Human cDNA encodin
c 23	47.4	2.3	1648	22	AAH26234	Human prothrombina
c 24	47.4	2.3	1651	22	AAS22580	Human cDNA encodin
c 25	47.4	2.3	1651	22	AAH26237	Human prothrombina
c 26	47.4	2.3	1824	22	AAD02607	Human angiopoietin
c 27	47.4	2.3	2013	22	AAH26235	Human prothrombina
c 28	47.4	2.3	2352	22	AAH26236	Human prothrombina
29	44.8	2.2	8318	20	AAX20264	Borrelia burgdorfe
c 30	44.4	2.1	4590	7	AAN60472	Sequence encoding
c 31	43.8	2.1	127197	22	AAI61370	Soybean 515002 reg
c 32	43	2.1	1218	21	AAA02488	Human colon cancer
c 33	43	2.1	11922	21	AAA70187	Plasmodium falcipa
c 34	42.6	2.1	114955	20	AAX53491	Human adenosine A1
35	42.2	2.0	381	22	AAH93355	Human chromosome 1
c 36	41	2.0	825	20	AAZ16536	Human gene express
37	41	2.0	1318	15	AAQ62614	Human mdr-1 promot
38	41	2.0	1688	15	AAQ62613	Human mdr-1 promot
c 39	41	2.0	1702	4	AAN30042	Sequence encoding
c 40	41	2.0	1702	19	AAV08180	Bovine interferon-
c 41	41	2.0	1702	20	AAV68163	Bovine interferon-
42	41	2.0	2090	15	AAQ62612	Human mdr-1 promot
43	40.6	2.0	580073	18	AAT58840	Mycoplasma genital
44	40.4	2.0	1337	20	AAZ17263	Human gene express
45	40.2	1.9	1231	20	AAX85025	Human secreted pro

ALIGNMENTS

RESULT 1
AAC85921
ID AAC85921 standard; DNA; 2069 BP.
XX
AC AAC85921;
XX
DT 22-AUG-2001 (first entry)
XX
DE MIP synthase promoter.
XX
KW Maize; myo-inositol-1 phosphate synthase; MIP synthase; promoter;
KW transgene; expression; embryo; transgenic plant; 5' UTR;
KW 5' untranslated region; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT promoter 7..2064
FT /*tag= a
FT /note= "Preferred promoter fragment"
XX
PN WO200140440-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32645.
XX
PR 02-DEC-1999; 99US-0168612.
XX
PA (DOWC) DOW AGROSCIENCES LLC.
XX
PI Armstrong K, Hey TD, Folkerts O, Smith KA, Hopkins NL;
XX
DR WPI; 2001-374815/39.

XX
PT Novel DNA molecules comprising maize myo-inositol-1 phosphate synthase
PT promoter, useful for controlling transgene expression in embryo tissue
PT of both monocotyledonous and dicotyledonous plants -
XX
PS Claim 2; Page 30-31; 31pp; English.
XX
CC This sequences represents the maize myo-inositol-1 phosphate (MIP)
CC synthase promoter. The promoter of maize MIP synthase is useful
CC for controlling transgene expression in embryo tissue of transgenic
CC plants, preferably maize plants. This sequence was isolated by
CC first isolating the maize MIP synthase gene, and then cloning the
CC associated 5' untranslated regions of the MIP synthase gene.
XX
SQ Sequence 2069 BP; 567 A; 531 C; 435 G; 536 T; 0 other;

Query Match 100.0%; Score 2069; DB 22; Length 2069;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2069; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 tctagatttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggag 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tctagatttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggag 60

Qy      61 aaattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaa 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aaattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaa 120

Qy     121 acctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtga 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 acctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtga 180

Qy     181 gttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaac 240

Qy     241 taagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctat 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 taagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctat 300

Qy     301 ttggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ttggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctc 360

Qy     361 acatcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 acatcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgc 420

Qy     421 tagtatgtgttataaaatcattttaccaattacctttttaaataactgtacgtagtttca 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 tagtatgtgttataaaatcattttaccaattacctttttaaataactgtacgtagtttca 480

Qy     481 tcagtagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagaga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tcagtagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagaga 540

Qy     541 tgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacga 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tgatgtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacga 600

Qy     601 caaaacaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattat 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 caaaacaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattat 660

Qy     661 tctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagta 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagta 720

Qy     721 aatattgctagcatttacgtctacggaatttattgaaaaaatgtagtattgttttatata 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 aatattgctagcatttacgtctacggaatttattgaaaaaatgtagtattgttttatata 780

Qy     781 attttaataaaactgtaaatcgtctggcttcgtttctggatggaggataaatagtgaata 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 attttaataaaactgtaaatcgtctggcttcgtttctggatggaggataaatagtgaata 840

Qy     841 cgaatgggaaacaccacaacaaccacgccgctgcgttctgcgaatcacatgagcgatcag 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 cgaatgggaaacaccacaacaaccacgccgctgcgttctgcgaatcacatgagcgatcag 900

Qy     901 tgccttgctgttccgtgaacttgcacgcaaggacgagagcctttctgcctttgcatgcaa 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 tgccttgctgttccgtgaacttgcacgcaaggacgagagcctttctgcctttgcatgcaa 960

Qy     961 ggacaagagtctttacatgcaaggacaaataactcccacgcgccccaccgtgctttggca 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ggacaagagtctttacatgcaaggacaaataactcccacgcgccccaccgtgctttggca 1020

Qy    1021 agccacatggcaccctgccgatcacaattcacaggcccaggcttccggtggtcgcgtgcc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 agccacatggcaccctgccgatcacaattcacaggcccaggcttccggtggtcgcgtgcc 1080

Qy    1081 gtgagtctgacaccgcaccacatggccgccgtaggccgtgcctacgcaccaaggcgactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 gtgagtctgacaccgcaccacatggccgccgtaggccgtgcctacgcaccaaggcgactc 1140

Qy    1141 gtggtgccaggctctcggcggctttggagtcggtgccatgccgcggggtccgtggaccgc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 gtggtgccaggctctcggcggctttggagtcggtgccatgccgcggggtccgtggaccgc 1200

Qy    1201 tcctaggggccaggacgaagctgcaccgcacaagcgggccgcgcgtgactccgtgaccgt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 tcctaggggccaggacgaagctgcaccgcacaagcgggccgcgcgtgactccgtgaccgt 1260

Qy    1261 gaggcgggcgtaaccaggagcttccgccacgcttgagaccacgtgacggcgcagaggagc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gaggcgggcgtaaccaggagcttccgccacgcttgagaccacgtgacggcgcagaggagc 1320

Qy    1321 tccacgcgatcaaaagcgcccgccacttctaaaggtcaggggtcttgcgttctgccccctc 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 tccacgcgatcaaaagcgcccgccacttctaaaggtcaggggtcttgcgttctgccccctc 1380

Qy    1381 gtgcttccttcaaattctggacctagtggatcaatttacgtacacctcagcaaccgatgc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 gtgcttccttcaaattctggacctagtggatcaatttacgtacacctcagcaaccgatgc 1440

Qy    1441 agccagtatgatgagcacgattgtgacgtgttgggggtcatggtcaatggcaaccgagca 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 agccagtatgatgagcacgattgtgacgtgttgggggtcatggtcaatggcaaccgagca 1500

Qy    1501 cgaattggtagtgtctgctttttgtacacgtgatagcatttgattcgttcattcaatttg 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 cgaattggtagtgtctgctttttgtacacgtgatagcatttgattcgttcattcaatttg 1560

Qy    1561 aactgtttaaacttatatatgtagagaaattagtccaactcatgcttaataaaaagtata 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 aactgtttaaacttatatatgtagagaaattagtccaactcatgcttaataaaaagtata 1620

Qy    1621 aaacccatcgaatttatgaattatgatagcaggtatcctatccattgtcatcgctcacag 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 aaacccatcgaatttatgaattatgatagcaggtatcctatccattgtcatcgctcacag 1680

Qy    1681 tcacagaggtagccactgccgacggccgacggcctcccatttcgctcccctcctactcct 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 tcacagaggtagccactgccgacggccgacggcctcccatttcgctcccctcctactcct 1740

Qy    1741 atgctgcggtccagcaaaagttcgggccttccggcaatccgccggcgcccgtcggctcaa 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 atgctgcggtccagcaaaagttcgggccttccggcaatccgccggcgcccgtcggctcaa 1800

Qy    1801 atcgcatctaccgcggctagaagctctctcttcctccctccgatccggtggggtccattt 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 atcgcatctaccgcggctagaagctctctcttcctccctccgatccggtggggtccattt 1860

Qy    1861 ccttcaattgtggcagtggccgtctcgaaccctctataaatcccccacccggacaccct 1920
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ccttcaattgtggcagtggccgtctcgaaccctctataaatccccccaccccggacaccct 1920

Qy    1921 tccccgaccacacagcccaacaacaaggagcgcggcggcccctccttcctttcctcccact 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 tccccgaccacacagcccaacaacaaggagcgcggcggcccctccttcctttcctcccact 1980

Qy    1981 tctctcgcgcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagt 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 tctctcgcgcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagt 2040

Qy    2041 gagaagggagggaattaaggcaaccatgg 2069
          |||||||||||||||||||||||||||||
Db    2041 gagaagggagggaattaaggcaaccatgg 2069


RESULT   2
AAF58252
ID   AAF58252 standard; DNA; 936 BP.
XX
AC   AAF58252;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1835.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;


  Query Match             8.9%;  Score 183.6;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 1.6e-37;
  Matches    6;  Conservative  533;  Mismatches  237;  Indels    0;  Gaps    0;

Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
           :::::::  ::: : ::: ::::::: :  : ::::::::    :::::::: ::  ::
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74

Qy      89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
           :::::::::: :   :: :::: :::::::::  ::::::::: ::: :: :   :: :
Db      75 wwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

Qy     149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
           ::: :::: :::::::: : :: ::::::: : : :: ::::: : :: ::: : :::::::
Db     135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

Qy     209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
           : ::: :::::::   :::::::::::   :: ::: :::  :  : : :::::: :::::
Db     195 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

Qy     269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
           : :  :: ::::: : : : ::: ::  :::::  : : :  ::::   ::: :: ::
Db     255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwww 314

Qy     329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
           :::::  ::    :  :::::::::  ::: : : ::   :  :  : :::::::: : ::
Db     315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

Qy     389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
           :: ::::  :::::::::|||     : : :  :: ::: : ::::::::: :::::::  :
Db     375 wwwwwwwwwwwwwwwwwgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 434

Qy     449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
           ::::  :::::::::: : ::  :: ::: :: : :: :: :: : :   :  :::::|
Db     435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 494

Qy     509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
           ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  ::::
Db     495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 554

Qy     569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
           :  :::::::::: :  :  : :::::  : :::: ::: :: :::: :::    :::
Db     555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwww 614

Qy     629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
           :::: :  : : : : :  : :  ::::::::: : : :  ::::: : :  :: ::: :
Db     615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 674

Qy     689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
           ::::::::::: ::: ::::: :::: :: :::::::  ::  :::::  : ::   ::
Db     675 wwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 734

Qy     749 tttattgaaaaaatgtagtattgttttatataattttaataaaactgtaaatcgtc 804
           :::::: :::::::: :: :::: :::::::::::::::::::::: : ::::: | |
Db     735 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcc 790


RESULT   3
AAF58254
ID   AAF58254 standard; DNA; 936 BP.
XX
AC   AAF58254;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1875.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
```

```
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

  Query Match             8.9%;  Score 183.6;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 1.6e-37;
  Matches    6;  Conservative  533;  Mismatches  237;  Indels    0;  Gaps    0;

Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74

Qy      89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
Db      75 wwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

Qy     149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
Db     135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

Qy     209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
Db     195 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

Qy     269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
Db     255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwww 314

Qy     329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
Db     315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

Qy     389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
Db     375 wwwwwwwwwwwwwwwwwgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 434

Qy     449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
Db     435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 494

Qy     509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
Db     495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 554

Qy     569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
Db     555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwww 614

Qy     629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
Db     615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 674

Qy     689 gtaaaaatttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
Db     675 wwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 734

Qy     749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaatcgtc 804
                                                                  | |
```

```
Db     735 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcc 790


RESULT   4
AAF58257
ID   AAF58257 standard; DNA; 936 BP.
XX
AC   AAF58257;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1954.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

  Query Match             8.9%;  Score 183.6;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 1.6e-37;
  Matches    6;  Conservative  533;  Mismatches  237;  Indels    0;  Gaps    0;

Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74

Qy      89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
Db      75 wwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

Qy     149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
Db     135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

Qy     209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
Db     195 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

Qy     269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
Db     255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwww 314
```

```
Qy     329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
           :::::  ::    :  ::::::::  ::: : : ::   :   :  : :::::::: : ::
Db     315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

Qy     389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
           :: ::::  ::::::::|||    : : :  :: ::: : ::::::::: :::::  :
Db     375 wwwwwwwwwwwwwwwwwgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 434

Qy     449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
           ::::  :::::::::::: : ::  :: ::: :: : :: :: :: : :   :  ::::::|
Db     435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 494

Qy     509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
           ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  :::::
Db     495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 554

Qy     569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
            :  :::::::::::: :  :  : :::::  : :::: ::: :: :::: :::    :::
Db     555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwww 614

Qy     629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
           :::: :  : : : : :  : :  ::::::::: : : :  ::::: : :  :: ::: :
Db     615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 674

Qy     689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
            :::::::::: ::: ::::: :::: :: :::::::  ::  :::::  : ::   ::
Db     675 wwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 734

Qy     749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaatcgtc 804
           :::::: ::::::: :: :::: :::::::::::::::::::: : ::::: | |
Db     735 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcc 790
```

RESULT 5
AAF58259
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 6; Page 128; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 8.9%; Score 183.6; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1.6e-37;
Matches 6; Conservative 533; Mismatches 237; Indels 0; Gaps 0;

```
Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
           :::::::  ::: : ::: ::::::: :  : :::::::::    :::::::: ::  ::
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74

Qy      89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
           ::::::::::  :   :: :::: :::::::::  :::::::::: ::: :: :   :: :
Db      75 wwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

Qy     149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
           ::: :::: :::::::: : :: :::::: : : :: ::::: : :: ::: : :::::::
Db     135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

Qy     209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
           : ::: ::::::   ::::::::::   :: ::: :::  :  : : :::::: ::::
Db     195 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

Qy     269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
           : :  :: ::::: : : : ::: ::  :::::  : : :  ::::   ::: :: ::
Db     255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwww 314

Qy     329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
           :::::  ::    :  ::::::::  ::: : : ::   :   :  : :::::::: : ::
Db     315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

Qy     389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
           :: ::::  ::::::::|||    : : :  :: ::: : ::::::::: :::::  :
Db     375 wwwwwwwwwwwwwwwwwgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 434

Qy     449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
           ::::  :::::::::::: : ::  :: ::: :: : :: :: :: : :   :  ::::::|
Db     435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 494

Qy     509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
           ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  :::::
Db     495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 554

Qy     569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
            :  :::::::::::: :  :  : :::::  : :::: ::: :: :::: :::    :::
Db     555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwww 614

Qy     629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
           :::: :  : : : : :  : :  ::::::::: : : :  ::::: : :  :: ::: :
Db     615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 674

Qy     689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
            :::::::::: ::: ::::: :::: :: :::::::  ::  :::::  : ::   ::
Db     675 wwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 734

Qy     749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaatcgtc 804
           :::::: ::::::: :: :::: :::::::::::::::::::: : ::::: | |
Db     735 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcc 790
```

RESULT 6
AAF58262
ID AAF58262 standard; DNA; 936 BP.
XX
AC AAF58262;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2007.
XX

```
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 128; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;
```

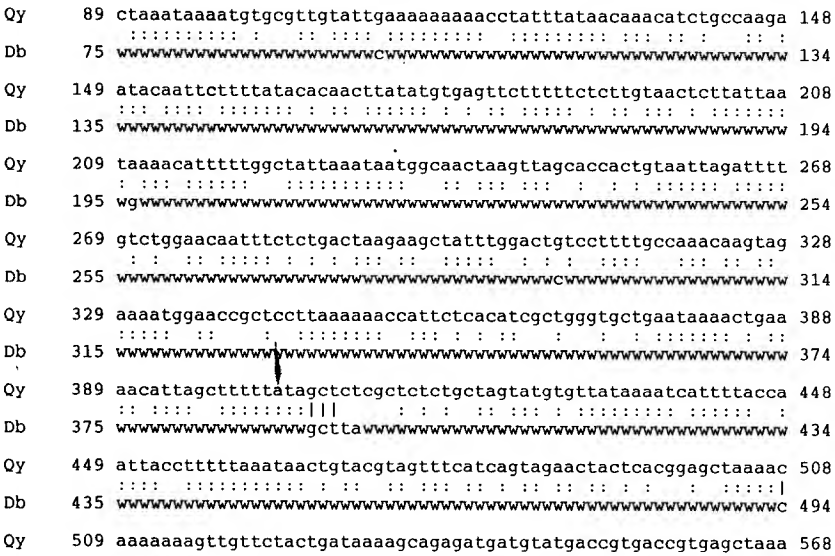
Query Match 8.9%; Score 183.6; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1.6e-37;
Matches 6; Conservative 533; Mismatches 237; Indels 0; Gaps 0;

```
Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
           :::::::  ::: : ::: :::::::: :  : :::::::::    :::::::: ::  ::
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74

Qy      89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
           :::::::::: :   :: :::: :::::::::  ::::::::: ::: :: : :   :: :
Db      75 wwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

Qy     149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
           ::: :::: :::::::: : :: :::::: : : :: ::::: : :: ::: : :::::::
Db     135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

Qy     209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
           : ::: :::::::   ::::::::::::   :: ::: :::  :  : : :::::: :::::
Db     195 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

Qy     269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
           : :  :: ::::: : : : ::: ::  :::::  : : :  ::::   ::: :: ::
Db     255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwww 314

Qy     329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
           :::::  ::    :  :::::::  ::: : : ::   :   :  : ::::::: : ::
Db     315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

Qy     389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
           :: ::::  ::::::::|||     : : :  :: ::: : ::::::::: :::::  :
Db     375 wwwwwwwwwwwwwwwwgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 434

Qy     449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
           ::::  :::::::::: : ::  :: ::: :: : :: :: :: : :   :  :::::|
Db     435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 494

Qy     509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
           ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  ::::
Db     495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 554

Qy     569 gtccaaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
           :  :::::::::::: :  :  : ::::: : :::: ::: :: :::: :::    :::
Db     555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwww 614

Qy     629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
           :::: :  : : : : :  : :  :::::::::: : : :  ::::: : :  :: ::: :
Db     615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 674

Qy     689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
           ::::::::::: ::: ::::: :::: :: ::::::::  ::  :::::  : ::   ::
Db     675 wwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 734

Qy     749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaatcgtc 804
           :::::: ::::::: :: :::: :::::::::::::::::::: : ::::: | |
Db     735 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcc 790
```

```
RESULT   7
AAF58255
ID   AAF58255 standard; DNA; 938 BP.
XX
AC   AAF58255;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1876.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
```

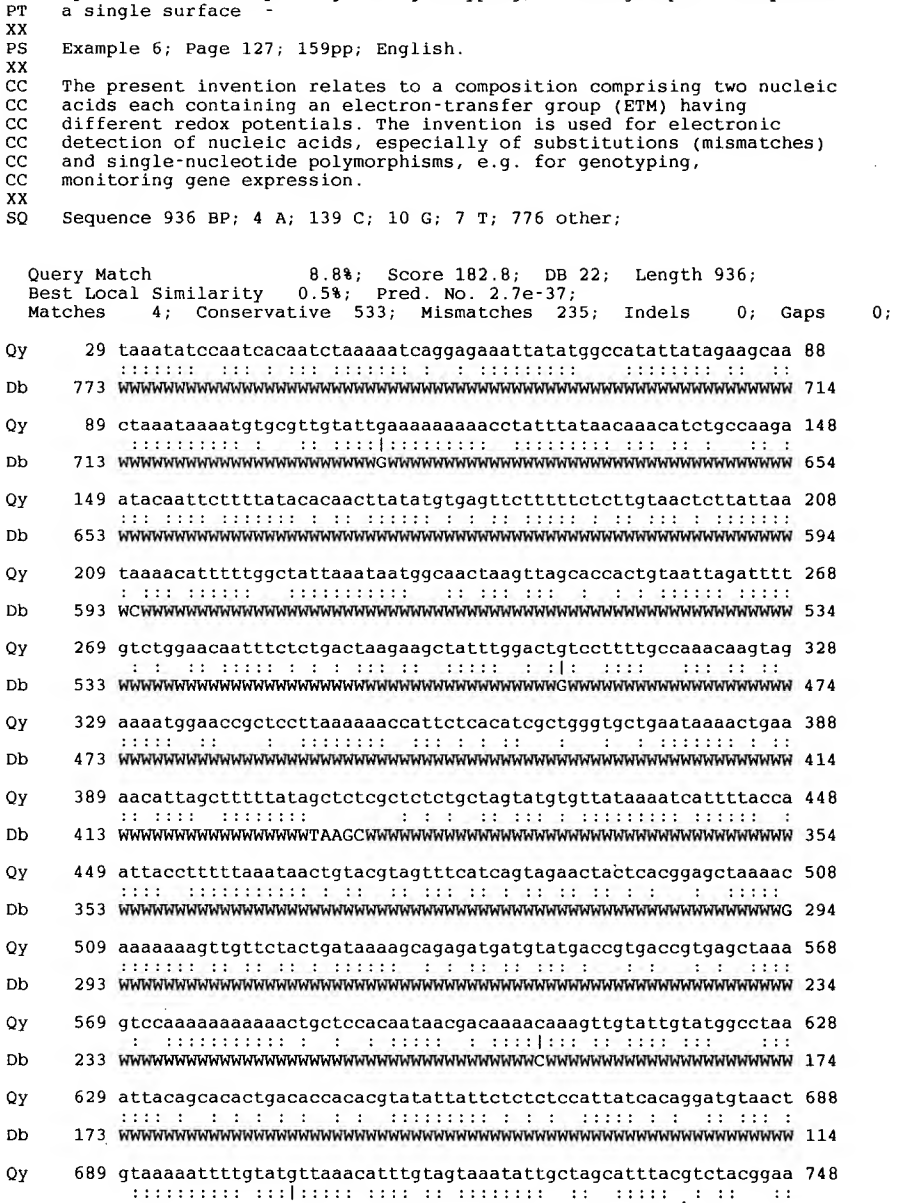
Query Match 8.9%; Score 183.6; DB 22; Length 938;
Best Local Similarity 0.8%; Pred. No. 1.7e-37;
Matches 6; Conservative 533; Mismatches 237; Indels 0; Gaps 0;

```
Qy      29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
           :::::::  ::: : ::: :::::::: :  : :::::::::    :::::::: ::  ::
Db      15 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 74
```



RESULT 8
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 8.8%; Score 182.8; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 2.7e-37;
Matches 4; Conservative 533; Mismatches 235; Indels 0; Gaps 0;



```
Db    113 WWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy    749 tttattgaaaaaatgtagtattgttttatataattttaataaaactgtaaat 800
          :::::: :::::::: :: :::: :::::::::::::::::::: : :::::
Db     53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2


RESULT   9
AAF58254/c
ID   AAF58254 standard; DNA; 936 BP.
XX
AC   AAF58254;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1875.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;


  Query Match             8.8%;  Score 182.8;  DB 22;  Length 936;
  Best Local Similarity   0.5%;  Pred. No. 2.7e-37;
  Matches    4;  Conservative  533;  Mismatches  235;  Indels    0;  Gaps    0;

Qy     29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
          :::::::  ::: : ::: ::::::: :  : ::::::::    :::::::: ::  ::
Db    773 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 714

Qy     89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
          :::::::::: :   :: ::::|:::::::::  ::::::::: ::: :: :   :: :
Db    713 WWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 654

Qy    149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
          ::: :::: ::::::: : :: :::::: : : :: ::::: : :: ::: : :::::::
Db    653 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 594

Qy    209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
          : ::: ::::::   :::::::::::   :: ::: :::  :  : : :::::: ::::
Db    593 WCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 534

Qy    269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
          : :  :: ::::: : : : ::: ::  :::::  : :|:  ::::   ::: :: ::
Db    533 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWW 474

Qy    329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
          :::::  ::    :  :::::::::  ::: : : ::   :   :  : :::::::: : ::
Db    473 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 414

Qy    389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
          :: ::::  :::::::::      : : :  :: ::: : ::::::::: :::::  :
Db    413 WWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 354

Qy    449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
          ::::  :::::::::::: : ::  :: ::: :: : :: :: :: : :   :  :::::
Db    353 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWG 294

Qy    509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
          ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  ::::
Db    293 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 234

Qy    569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
          :  :::::::::::: :  :  : :::::  : ::::|::: :: :::: :::    :::
Db    233 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWW 174

Qy    629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
          :::: :  : : : :  :  :  ::::::::: : : :  :::: : :  :: ::: :
Db    173 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 114

Qy    689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
          :::::::::: :::|::::: :::: :: :::::::  ::  :::::  : ::   ::
Db    113 WWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy    749 tttattgaaaaaatgtagtattgttttatataattttaataaaactgtaaat 800
          :::::: :::::::: :: :::: :::::::::::::::::::::: : :::::
Db     53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2


RESULT  10
AAF58257/c
ID   AAF58257 standard; DNA; 936 BP.
XX
AC   AAF58257;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1954.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
```

CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 8.8%; Score 182.8; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 2.7e-37;
Matches 4; Conservative 533; Mismatches 235; Indels 0; Gaps 0;

```
Qy     29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
Db    773 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 714

Qy     89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
Db    713 WWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 654

Qy    149 atacaattcttttatacacaacttatatgtgagttcttttttctcttgtaactcttattaa 208
Db    653 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 594

Qy    209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
Db    593 WCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 534

Qy    269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
Db    533 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWW 474

Qy    329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
Db    473 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 414

Qy    389 aacattagcttttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
Db    413 WWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 354

Qy    449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
Db    353 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWG 294

Qy    509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
Db    293 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 234

Qy    569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
Db    233 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWW 174

Qy    629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
Db    173 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 114

Qy    689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
Db    113 WWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy    749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaat 800
Db     53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2
```

RESULT 11
AAF58259/c
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 128; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 8.8%; Score 182.8; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 2.7e-37;
Matches 4; Conservative 533; Mismatches 235; Indels 0; Gaps 0;

```
Qy     29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
Db    773 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 714

Qy     89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
Db    713 WWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 654

Qy    149 atacaattcttttatacacaacttatatgtgagttcttttttctcttgtaactcttattaa 208
Db    653 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 594

Qy    209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
Db    593 WCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 534

Qy    269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
Db    533 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWW 474

Qy    329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
Db    473 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 414

Qy    389 aacattagcttttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
Db    413 WWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 354

Qy    449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
```

```
               ::::  ::::::::::: : ::  :: ::: :: : :: :: :: : :   :  :::::
Db       353 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWG 294

Qy       509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
             :::::: :: :: :: : ::::::  : : :: :: ::: :    : :    : :  ::::
Db       293 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 234

Qy       569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
              :  :::::::::: :   :   : :::::  : ::::|::: :: :::: :::     :::
Db       233 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWW 174

Qy       629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
              :::: :  : : : : :  : :  :::::::: : : :  ::::: : :  :: ::: :
Db       173 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 114

Qy       689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
               :::::::::: :::|::::: :::: :: ::::::::  ::  :::::  : ::   ::
Db       113 WWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy       749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaat 800
             :::::: ::::::: :: :::: ::::::::::::::::::: : :::::
Db        53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2
```

RESULT 12
AAF58262/c
ID   AAF58262 standard; DNA; 936 BP.
XX
AC   AAF58262;
XX
DT   24-APR-2001 (first entry)
XX
DE   Oligonucleotide D2007.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface -
XX
PS   Example 6; Page 128; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

```
  Query Match             8.8%;  Score 182.8;  DB 22;  Length 936;
  Best Local Similarity   0.5%;  Pred. No. 2.7e-37;
  Matches    4;  Conservative  533;  Mismatches  235;  Indels    0;  Gaps    0;

Qy        29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
              ::::::  ::: : ::: ::::::: :  : :::::::::     :::::::: ::  ::
Db       773 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 714

Qy        89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaga 148
              ::::::::: :     :: ::::|:::::::::  :::::::: ::: :: :    :: :
Db       713 WWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 654

Qy       149 atacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaa 208
             ::: :::: ::::::: : :: :::::: : : :: ::::: : :: ::: : :::::::
Db       653 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 594

Qy       209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
              : ::: ::::::     ::::::::::     :: ::: :::    :    : : :::::: :::::
Db       593 WCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 534

Qy       269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
              : :  :: ::::: : : : ::: ::  :::::    : :|:  ::::    ::: :: ::
Db       533 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWW 474

Qy       329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
              :::::  ::     :  ::::::::  ::: : : ::     :    :  : ::::::: : ::
Db       473 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 414

Qy       389 aacattagcttttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
             :: ::::  ::::::::          : : :   :: ::: : ::::::::: ::::::   :
Db       413 WWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 354

Qy       449 attaccttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
             ::::  ::::::::::: : ::  :: ::: :: : :: :: :: : :     :  :::::
Db       353 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWG 294

Qy       509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
             :::::: :: :: :: : ::::::  : : :: :: ::: :    : :    : :  ::::
Db       293 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 234

Qy       569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
              :  :::::::::: :   :   : :::::  : ::::|::: :: :::: :::     :::
Db       233 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWW 174

Qy       629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
              :::: :  : : : : :  : :  :::::::: : : :  ::::: : :  :: ::: :
Db       173 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 114

Qy       689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
               :::::::::: :::|::::: :::: :: ::::::::  ::  :::::  : ::   ::
Db       113 WWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy       749 tttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaat 800
             :::::: ::::::: :: :::: ::::::::::::::::::: : :::::
Db        53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2
```

RESULT 13
AAF58255/c
ID   AAF58255 standard; DNA; 938 BP.
XX
AC   AAF58255;
XX
DT   24-APR-2001 (first entry)
XX
DE   Oligonucleotide D1876.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX

```
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
```

Query Match 8.8%; Score 182.8; DB 22; Length 938;
Best Local Similarity 0.5%; Pred. No. 2.7e-37;
Matches 4; Conservative 533; Mismatches 235; Indels 0; Gaps 0;

```
Qy     29 taaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaa 88
          :::::::  ::: : ::: :::::::: :  : :::::::::    :::::::: ::  ::
Db    773 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 714

Qy     89 ctaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaaaga 148
           :::::::::: :   :: ::::|:::::::::  :::::::::: ::: :: :   :: :
Db    713 WWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 654

Qy    149 atacaattcttttatacacaacttatatgtgagttcttttttctcttgtaactcttattaa 208
          ::: :::: ::::::: : :: :::::: : : :: ::::: : :: ::: : :::::::
Db    653 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 594

Qy    209 taaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagatttt 268
          : ::: ::::::   :::::::::::   :: ::: :::  :  : : :::::: :::::
Db    593 WCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 534

Qy    269 gtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtag 328
           : :  :: ::::: : : : ::: ::  :::::  : :|:  ::::   ::: :: ::
Db    533 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWW 474

Qy    329 aaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaa 388
          :::::  ::    :  :::::::::  ::: : : ::   :   :  : ::::::: : ::
Db    473 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 414

Qy    389 aacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttacca 448
          :: ::::  ::::::::        : : :  :: ::: : ::::::::: ::::::  :
Db    413 WWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 354

Qy    449 attacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaac 508
          ::::  :::::::::::: : ::  :: ::: :: : :: :: :: : :   :  :::::
Db    353 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWG 294

Qy    509 aaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaa 568
          ::::::: :: :: :: : ::::::  : : :: :: ::: :   : :   : :  :::
Db    293 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 234

Qy    569 gtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaa 628
           :  :::::::::: :  :  : ::::  : ::::|::: :: :::: :::    :::
Db    233 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWW 174

Qy    629 attacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaact 688
          :::: :  : : : : :  : :  ::::::::: : : :  ::::: : :  :: ::: :
```

```
Db    173 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 114

Qy    689 gtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaa 748
          :::::::::: :::|::::: :::: :: :::::::  ::  :::::  : ::   ::
Db    113 WWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 54

Qy    749 tttattgaaaaaatgtagtattgttttatataattttaataaaactgtaaat 800
          :::::: ::::::: :: :::: :::::::::::::::::::::: : ::::::
Db     53 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 2
```

```
RESULT  14
AAC35815
ID   AAC35815 standard; DNA; 483 BP.
XX
AC   AAC35815;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Zea mays DNA fragment SEQ ID NO: 11525.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic;
KW   pathway; promoter; termination sequence; corn; ss.
XX
OS   Zea mays subsp. mays.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
```

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.9%; Score 122.2; DB 21; Length 483;
Best Local Similarity 95.8%; Pred. No. 8.4e-22;
Matches 136; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 1922 ccccgaccacacagcccaacaacaaggagcgcggcggcccctccttccttcctcccactt 1981

```
     | | | ||||||||||||||||  |||||||||||||  ||||||||||||||||||||||
Db     2 cacggtccacacagcccaaca--aaggagcgcggcgnccccctccttccttcctcccactt 59

Qy  1982 ctctcgcgcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagtg 2041
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    60 ctctcgcgcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagtg 119

Qy  2042 agaagggagggaattaaggcaa 2063
         |||||||||||||||||||||
Db   120 agaagggagggaattaaggcaa 141


RESULT  15
AAC85922
ID   AAC85922 standard; cDNA; 1959 BP.
XX
AC   AAC85922;
XX
DT   22-AUG-2001  (first entry)
XX
DE   MIP synthase cDNA.
XX
KW   Maize; myo-inositol-1 phosphate synthase; MIP synthase; promoter;
KW   transgene; expression; embryo; transgenic plant; 5' UTR;
KW   5' untranslated region; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             137..1669
FT                   /*tag= a
FT                   /product= "MIP synthase"
XX
PN   WO200140440-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32645.
XX
PR   02-DEC-1999;   99US-0168612.
XX
PA   (DOWC ) DOW AGROSCIENCES LLC.
XX
PI   Armstrong K,  Hey TD,  Folkerts O,  Smith KA,  Hopkins NL;
XX
DR   WPI; 2001-374815/39.
DR   P-PSDB; AAB47286.
XX
PT   Novel DNA molecules comprising maize myo-inositol-1 phosphate synthase
PT   promoter, useful for controlling transgene expression in embryo tissue
PT   of both monocotyledonous and dicotyledonous plants -
XX
PS   Example 1; Page 26-28; 31pp; English.
XX
CC   This sequences encodes maize myo-inositol-1 phosphate (MIP)
CC   synthase. This sequence was used in the isolation of the MIP synthase
CC   promoter sequence.  The promoter of maize MIP synthase is useful
CC   for controlling transgene expression in embryo tissue of transgenic
CC   plants, preferably maize plants.  The promoter sequence was isolated
CC   by first isolating the maize MIP synthase gene; and then cloning the
CC   associated 5' untranslated regions of the MIP synthase gene.
XX
SQ   Sequence 1959 BP; 472 A; 518 C; 568 G; 401 T; 0 other;


  Query Match             5.8%;  Score 119;  DB 22;  Length 1959;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-20;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1945 aaggagcgcggcggcccctccttccttcctcccacttctctcgcgcggcgctcgcttacc 2004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17 aaggagcgcggcggcccctccttccttcctcccacttctctcgcgcggcgctcgcttacc 76

Qy   2005 tcgcctcgcattccgttcgagcaggggagcggcagtgagaagggagggaattaaggcaa 2063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 tcgcctcgcattccgttcgagcaggggagcggcagtgagaagggagggaattaaggcaa 135


Search completed: December 19, 2001, 22:42:41
Job time: 4731 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 18:34:04 ; Search time 1766.62 Seconds
(without alignments)
12585.058 Million cell updates/sec

Title: US-09-727-628-3
Perfect score: 2069
Sequence: 1 tctagatttttttttcaattc.....gggaattaaggcaaccatgg 2069

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	71	3.4	1101	13	CNS0039G	AL063921 Drosophil
2	70.2	3.4	1101	13	CNS0039G	AL063921 Drosophil
3	67	3.2	582	10	AI770934	AI770934 606063D12
c 4	64	3.1	1101	13	CNS00EO7	AL069440 Drosophil
5	58.8	2.8	1101	13	CNS00EO7	AL069440 Drosophil
6	58.4	2.8	1101	13	CNS00EVL	AL069706 Drosophil
7	58	2.8	454	10	AW231748	AW231748 687023G08
c 8	58	2.8	1201	13	CNS0167M	AL106396 Drosophil
9	57.4	2.8	1101	13	CNS0021J	AL061936 Drosophil
c 10	56.4	2.7	928	13	CNS00DKY	AL071865 Drosophil
11	56	2.7	925	13	CNS0091P	AL053013 Drosophil
12	54.8	2.6	500	10	AU086608	AU086608 AU086608
13	54.4	2.6	869	13	CNS017Z2	AL108680 Drosophil
c 14	53.8	2.6	925	13	CNS0091P	AL053013 Drosophil
c 15	53.4	2.6	987	13	CNS014PQ	AL104456 Drosophil
16	53.2	2.6	902	11	BG321216	BG321216 Zm04_05g1
c 17	53	2.6	1248	13	B11336	B11336 F19M10-Sp6
18	52.8	2.6	1101	13	CNS00BO1	AL057419 Drosophil
c 19	52.2	2.5	1001	13	CNS0140O	AL103554 Drosophil
c 20	52.2	2.5	1101	13	CNS00YWL	AL096927 Drosophil
21	52.2	2.5	1200	13	CNS016CO	AL106578 Drosophil
c 22	52.2	2.5	1225	13	CNS0161D	AL106171 Drosophil
23	51.6	2.5	975	13	CNS0102Y	AL098452 Drosophil
24	51.4	2.5	867	13	CNS075BG	AL429890 clone BA0
25	51.2	2.5	912	13	CNS006N3	AL065775 Drosophil
26	51.2	2.5	994	13	CNS015XG	AL106030 Drosophil
c 27	51.2	2.5	1101	13	CNS000D1	AL065414 Drosophil
28	51.2	2.5	1101	13	CNS00LVZ	AL078819 Drosophil
29	51.2	2.5	1101	13	CNS00ZBM	AL097468 Drosophil
30	51	2.5	1099	10	AL536986	AL536986 AL536986
c 31	51	2.5	1101	13	CNS00EVL	AL069706 Drosophil
c 32	50.8	2.5	1051	10	AL535414	AL535414 AL535414
c 33	50.8	2.5	1101	13	CNS00Z2U	AL097152 Drosophil
c 34	50.6	2.4	500	10	AU087038	AU087038 AU087038
c 35	50.6	2.4	976	13	CNS04E5M	AL286627 Tetraodon
c 36	50.4	2.4	713	13	CNS06CIJ	AL392561 T7 end of
37	50.4	2.4	928	13	CNS00DKY	AL071865 Drosophil
38	50.4	2.4	1101	13	CNS00YWL	AL096927 Drosophil
39	50.4	2.4	1364	11	BF033991	BF033991 601456301
40	50.2	2.4	1101	13	CNS00DT7	AL075293 Drosophil
c 41	50	2.4	942	13	CNS018GS	AL109318 Drosophil
42	50	2.4	1101	13	CNS017SY	AL108460 Drosophil
43	50	2.4	1225	13	CNS0161D	AL106171 Drosophil
c 44	49.8	2.4	631	10	AL566993	AL566993 AL566993
45	49.8	2.4	749	13	AQ794523	AQ794523 nbxb0053M

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS CNS0039G 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR08K10"
                     /note="end : TET3"
BASE COUNT      201 a     64 c    131 g    202 t    503 others
ORIGIN
```

```
  Query Match             3.4%;  Score 71;  DB 13;  Length 1101;
  Best Local Similarity   17.3%;  Pred. No. 5.4e-05;
  Matches  103;  Conservative  254;  Mismatches  240;  Indels    0;  Gaps    0;

Qy     204 attaataaaacatttttggctattaaataatggcaactaagttagcaccactgtaattag 263
           ::|:: : |:  :::|:   :::::|::::: | |: :|::: || :  | : ::::::
Db    1094 DDTWDRDTRKDDWDTTKWWTWWKDRADDRRWAGDADRWAWDDGAGTWWTATWWWWWWWWA 1035

Qy     264 attttgtctggaacaatttctctgactaagaagctatttggactgtccttttgccaaaca 323
           ::|: : |:  :| |:|:|   | : ||::|:   ::   |  :::  |: :|  :::
Db    1034 TWDTWWDKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKRDRDAATDADGAGRRDGG 975

Qy     324 agtagaaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaa 383
           ::  :  :|:::|:::  :    ::|| ||    :: : : :: : : :| :  :|:::: |
Db     974 RKRKDKKDRKDGDDDKKGGKKKKAAKAAKWATKWWDDWDWDKDWKWDGAKDRKADDDDGA 915

Qy     384 ctgaaaacattagctttttatagctctcgctctctgctagtatgtgttataaaatcattt 443
           :::::   ::|: :|: | ::: : : : : : :| : :  :  :  |::::: : : ::
Db     914 GDKDDDGKGKDADDDTDGTKDDDDKDKWDDWDKAKGTWGDATWAWAATDWWWWGWADADW 855

Qy     444 taccaattaccttttaaataactgtacgtagtttcatcagtagaactactcacggagct 503
           :   || :   ::::|: :  :: :   |:| : ::  : || :  | : : :::  :
Db     854 WTWDAAADDWWADDWDAWAWKWDDAWAWGARTADRRDWGDRAGKRGGARKRRDRKRADD 795

Qy     504 aaaacaaaaaaagtttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgag 563
           ::| :::::||  :  :| |   : ::: :: : : :: |::|:::: ::::  ||  :
Db     794 KRDAADDRDDAATWTTWTTTTTRDTDDWKWKTDTWTRWAADRTWDRDDDDDRDRAGTAGR 735

Qy     564 ctaaagtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatgg 623
           :::  : |: :::::::|  : |  : :: :  :  :|      | :: : : ::  :::
Db     734 KWRRTWKRRWKRRDTRWDDADADDTARDDRRRRGDDGADAGKGKKTGRKRRRRDRATWDR 675

Qy     624 cctaaattacagcaactgacaccacacgtatattattctctctccattatcacaggatg 683
              ::|: | :      |:: :  : : : | :  :||     :   ::|: :   :::
Db     674 TDAWWADAAWWTTTDTDTDDWDKRDRRRKGARRRRRTTARAAWDWWTWKAWDWAKWDWKT 615

Qy     684 taactgtaaaaatttgtgtatgttaaacatttgtagtaaatattgctagcatttacgtcta 743
           |: :: :||: :||  ::|::::| | |:  : :: :|| |: : | :  ||||  | :|
Db     614 RADRWDRWAADTWTDARKADRDWAKARAWRARRDRARAARADRRWTTKGKTTTATWTTWA 555

Qy     744 cggaatttattgaaaaaatgtagtattgttttatataattttaataaaactgtaaat 800
           : |:: :|: :   | :| |: | :| |||| | | :||| |:::||: | |:||:
Db     554 ARAAWWAWWAWATTTATWTTTWTTTWTTTTTTTTTTTWTTTAAWWWAAWWTATWAAW 498
```

```
RESULT   2
CNS0039G
LOCUS       CNS0039G     1101 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR08K10"
                     /note="end : TET3"
BASE COUNT      201 a     64 c    131 g    202 t    503 others
ORIGIN
```

```
  Query Match             3.4%;  Score 70.2;  DB 13;  Length 1101;
  Best Local Similarity   18.0%;  Pred. No. 8e-05;
  Matches  123;  Conservative  271;  Mismatches  289;  Indels    0;  Gaps    0;

Qy      90 taaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaa 149
           :|||:|||||     :  :|:  ||||  ::|   :   |:|: |::  | :   || ||
Db     417 WAAAWAAAATAATTWWAAWAWAAAAAATTWWAAAAWAAAAWAWTAWWTTTAWTWAAAAAA 476

Qy     150 tacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaat 209
           | || | :||| |  :    :| :|| :   :: ||          ||   : | :||
Db     477 AAAAAATTWTTTTTTTWTTTAWTTWATAWWTTWWWTTAAAWAAAAAAAAAAAWAAAWAAA 536

Qy     210 aaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagattttg 269
           :| | ||:|::   | :| :| :||   : : |: ::    :  :: | :|:   ::::
Db     537 WATAAATWTWWTWWTTYTTWAAWATAAAMCMAAWYYHTYTTYTYHYYTYWTYTMTWHYHT 596

Qy     270 tctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaagtaga 329
           :|  :|:  ::: ::| |:::  :  |:: :  : | |  :::  |:  ::  :: :
Db     597 MYTHAWAHTTWYHWYHTYAMWHWMTWHWTMWAWWHWTTYTAAYYYYYTCMYYYHYMHWHH 656

Qy     330 aaatggaaccgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaaa 389
           |:|:  |:   :|  |:| ::| :: ::: :  : | ||   : |:   ::  |:: : :
Db     657 AHAHAAAWWTTHTWWTHAYHWATYHYYYYMYCAMMCMCTHTCHHCYYYYHHYTAHHTHTH 716

Qy     390 acattagctttttatagctctcgctctctgctagtatgtgttataaaatcattttaccaa 449
           :   :  :::: :|:  ::||   ::::: ::  :|: | ::|:|:|:::::: :   ||
Db     717 HWYAHYYMWYYMWAYYWMYCTACTYHYHHHHHYHWAYHTTWYAWAHAMWMWHHAHYAAAA 776

Qy     450 ttacctttttaaataactgtacgtagtttcatcagtagaactactcacggagctaaaaca 509
           :|   ||:: ::| ::  : : :: : :   :: |   ::: :|       ::::: :
Db     777 AWAAWATTHHYHHTTHYMHHTYMYHYYMYTCCYMCTYHCWHYYHTAYTCWTWTHHWMWTW 836

Qy     510 aaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaag 569
           :: :: :: :| |:   :: : :  : : || :|   : : | | : : : : :::::
Db     837 THWYHHTWWHHTTTHWAWWHTHTWCWWWWHATTWTWATHCWACMTMHWHHWMHMHHHHMA 896

Qy     570 tccaaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcctaaa 629
           : ::: :: : :::  |||::::   ::  :::::::::: :: :: |   |   : ::
Db     897 CHAHHHTHMCMCHHHMHCTCHHHHTMYHMTCHWWWHMHWHWHHWWMATWMTTMTTMMMMC 956

Qy     630 ttacagcacactgacaccacacgtatattattctctctccattatcacaggatgtaactg 689
             ::: |:: :  :: :: |::    :|:|||:::::  | :: | :   |: :|:
Db     957 CMMHHHCHMYHMMHMYMYCCHYYCTCHTHATTHYHYMCTCYHYCTWHTYWTAYWWAWTAH 1016
```

```
Qy     690 taaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacggaat 749
           :  | :::: ::: ::| : :::: :|  ::|   :  |: : : |:: ::| :  :|:
Db    1017 AMTTATWWWWMHWWAHWATWWWWWWWWATAWWACTCHHWTWYHTHCTWYYHHTYHMWWAW 1076

Qy     750 ttattgaaaaaatgtagtattgt 772
           : |:: ::: |:: :| : :: |
Db    1077 WMAWHWHHMYAHYHWAHHCWYYT 1099


RESULT   3
AI770934
LOCUS       AI770934      582 bp    mRNA            EST       02-FEB-2000
DEFINITION  606063D12.x1 606 - Ear tissue cDNA library from Schmidt lab Zea
            mays cDNA, mRNA sequence.
ACCESSION   AI770934
VERSION     AI770934.1  GI:5268970
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 606063  row: D  column: 12.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Zea mays"
                     /cultivar="Ohio43"
                     /db_xref="taxon:4577"
                     /clone_lib="606 - Ear tissue cDNA library from Schmidt
                     lab"
                     /tissue_type="mixed"
                     /dev_stage="ear length from 0.5 cm - 2.0 cm"
                     /lab_host="XLOLR (Stratagene)"
                     /note="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
                     ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
                     lab"
BASE COUNT      112 a    199 c    169 g    102 t
ORIGIN


  Query Match             3.2%;  Score 67;  DB 10;  Length 582;
  Best Local Similarity   66.1%;  Pred. No. 0.00038;
  Matches  170;  Conservative    0;  Mismatches   60;  Indels   27;  Gaps    4;

Qy    1802 tcgcatctaccgcggctagaagctctctcttcc--tccctccgatccggtggggtccatt 1859
           | ||  |  |||||||| |  || ||| |||  ||||| ||||||||||||||||| ||
Db       2 TTGCCCCTGACGCGGCTACGACGTCGCTCCTCCGGTCCCTTCGATCCGGTGGGGTCCGTT 61

Qy    1860 tccttcaattgtggcagtggccgtc----------------tcgaaccctctataaat 1901
           || ||||  | |||   |   | |                ||||||||||||||||
Db      62 TCTTTCAAGCGCGGCCTCGCTGGCCGCCTCGTGGCAGTGACCGTCGAACCCTCTATAAAT 121

Qy    1902 cccccaccccggacacccttccccgaccacacagcccaacaacaaggagcgcggcggccc 1961
           |||  |||||  |||||||||| ||| |||||| |||||      |   |   |   |
Db     122 CCCGTGCCCCGAGCACCCTTCCTCGATCACACAACCCAAA------GCAGCCACAGCAGC 175

Qy    1962 ctccttcctcctcccacttctctcgcgcggcgctcgcttacctcgcctcgcattccgtt 2021
           |||||||| |||  |   || |||||||| |||||||| |||||||| |||||||| ||
Db     176 CTCCTTCCTTCTCTCACTCTCGCTCGCGCTGCGCTCGC-TACCTCGCTTCGCATTCCATT 234

Qy    2022 cgagcaggggagcggca 2038


          |||  || || | || |
Db     235 CGAAAAGAGGGGAGGAA 251


RESULT   4
CNS00EO7/c
LOCUS       CNS00EO7     1101 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069440
VERSION     AL069440.1  GI:4949583
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR29P01"
                     /note="end : TET3"
BASE COUNT      366 a     66 c    104 g    351 t    214 others
ORIGIN


  Query Match             3.1%;  Score 64;  DB 13;  Length 1101;
  Best Local Similarity   37.6%;  Pred. No. 0.0017;
  Matches  172;  Conservative   62;  Mismatches  219;  Indels    4;  Gaps    1;

Qy       2 ctagatttttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggaga 61
           | : |: |||::|  |   |       |  |||| ||||   ||  ||||||: |  : |
Db     973 CAWWAYATTTWWTATACATAATWWTTTTTTATATACAATTTWAAAATAAAAAWTAACWWA 914

Qy      62 aattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaa 121
           | |||:|   |  |:|| |  |:  || | :|| | |:     :|  ::| |||||| :
Db     913 ATTTAWAAAWCATTWTTTTTAAWWTAATTTWATTACAW----TWTTAWWTAAAAAAATWT 858

Qy     122 cctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtgag 181
             |:| | :| ||| :  |   |   |:| :||  :|  |:: | |: |||   : |
Db     857 TAAAWTAAAWAAAAARWATTAWAATTTAYATWATATWTAAAWWTATAWATTATTMWWAAT 798

Qy     182 ttctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaact 241
           :|  | :  | || |   | : |:|: |:| : :|  ||    :|::   |||  ||
Db     797 WTTATAWNATTTTTTTTWTAWAAWTWTTWATWAWTAATTTTAAWTWWTTAAATAAAAAAA 738

Qy     242 aagttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctatt 301
           |  |         | | |||: | :|| | |   : | |:::| | |  :  ||  |:||
Db     737 ATTTATTTTTATTTWTTATTWAAAWTTTTWTTTTWAATTWWYTTTAATAWTTAAAWTWTT 678
```

```
Qy    302 tggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctca 361
          : : |  :||    ||| :  |  |||||  :|    |   |  |:|:    | :
Db    677 AAAWAWWTAAWTTAAATAAATWTATTTAAAATWTWAAWTATAAATTTAWADWTTATAWWT 618

Qy    362 catcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgct 421
          |   :   |     ||| ||   :||| ||||  ||:|| |   :   |     :
Db    617 TTTTTTWWCWTTWAATATATAAAATWAAAAATTAAATTWTTTTTATATWWWATAAAAAMA 558

Qy    422 agtatgtgttataaaatcattttaccaattacctttt 458
            ::|   ||||||::   | ||   |:||   ||||
Db    557 TWWWTTATTTATAAWWAATTATTTAWAWTTGATTTTT 521
```

RESULT 5
CNS00EO7
LOCUS CNS00EO7 1101 bp DNA GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069440
VERSION AL069440.1 GI:4949583
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR29P01"
/note="end : TET3"
BASE COUNT 366 a 66 c 104 g 351 t 214 others
ORIGIN

Query Match 2.8%; Score 58.8; DB 13; Length 1101;
Best Local Similarity 35.4%; Pred. No. 0.022;
Matches 219; Conservative 75; Mismatches 324; Indels 1; Gaps 1;

```
Qy     16 aattcaccccgagtaaatatccaatcacaatctaaaaatcaggagaaattatatggccat 75
          ||   | |  :| :||| :    |: |  |: |  |||     | |||  | ||    |
Db    419 AAAAAAACATKAAWAAARRGGGGAKAAATAWTTKGAAAGATATAAAAAAAARATAAWTAA 478

Qy     76 attatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaa 135
          : |  | ||  :|   |||||||: |       :    ||| ||||   | :|:|||  |
Db    479 WATTAATAAATWATATAATAAAAWTTTAAAAAAWAAATAAATAAAAATCAAWTWTAAATA 538

Qy    136 acatctgccaagaatacaattcttttatacacaacttatatgtgagttctttttctcttg 195
          |  : |   ||  | : :||: ||||||: :  |    | |:        ||||: | ||
Db    539 ATTWWTTATAAATAAWWWATKTTTTTATWWWATATAAAAAWAATTTAATTTTTWATTTTA 598

Qy    196 taactcttattaataaaacatttttggctattaaataatggcaactaagttagcaccact 255
          ||  | : |: :: |||| |::| |   : :|||||          :| ||   |  : |
Db    599 TATATTWAAWGWWAAAAAAAWWTATAAWHTWTAAATTTATAWTTWAWATTTTAAATAWAT 658

Qy    256 gtaattagattttgtctggaacaatttctctgactaagaagctatttggactgtcctttt 315
          || ||| :|| : | |  || |:|||   | | ||| ::    |:   |     |||:
Db    659 TTATTTAAWTTAWWTWTTTAAWAWTTTAAWT-ATTAAARWWAATTWAAAAWAAAAWTTTW 717

Qy    316 gccaaacaagtagaaaatggaaccgctccttaaaaaaccattctcacatcgctgggtgct 375
             ||: ||  | || |         |  ||||::|:  |   | |  :       :  :
Db    718 AATAAWAAATAAAAATAAATTTTTTTTATTTAAWWAWTTAAAATTAWTWATWAAWAWTTW 777

Qy    376 gaataaaactgaaaacattagctttttatagctctcgctctctgctagtatgtgttataa 435
          |: ||||     :     :  |::: |||  :     : |    || :||:|    |:
Db    778 TAWAAAAAAAATNWTATAAWATTWWKAATAATWTATAWWTTTAWATATWATRTAAATTWT 837

Qy    436 aatcattttaccaattacctttttaaataactgtacgtagtttcatcagtagaactactc 495
          |||  ||||      :     : |   |   : ||   :  |   : | |  :: ||
Db    838 AATWYTTTTTWTTTAWTTTAAWATTTTTTTAWWTAAWAWTGTAATWAAATTAWWTTAAAA 897

Qy    496 acggagctaaaacaaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtg 555
          |    | |  :  ||  ::  |  || |  |  |:|||    | || |   |: |   ||
Db    898 AWAATGWTTTWTAAATTWWGTTAWTTTTTATTTTWAAATTGTATATAAAAAAWWATTATG 957

Qy    556 accgtgagctaaagtccaaaaaaaaaaaactgctccacaataacgacaaaacaaagttgta 615
             : |  |: : |  |   |||| |   : |    || ||  | | |: ||| : :|
Db    958 TATAWWAAATRTWWTGTATTKAAAATATWGRATATTKAAAAAAAAAATAWTAAAAWARTK 1017

Qy    616 ttgtatggcctaaattaca 634
          :|||:|     |||  | |
Db   1018 KTGTRTWTKWAAAAAAAAA 1036
```

RESULT 6
CNS00EVL
LOCUS CNS00EVL 1101 bp DNA GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1101

```
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR29B23"
                     /note="end : T7"
BASE COUNT      419 a     91 c     60 g    299 t    232 others
ORIGIN


  Query Match             2.8%;  Score 58.4;  DB 13;  Length 1101;
  Best Local Similarity   30.2%;  Pred. No. 0.027;
  Matches  186;  Conservative  109;  Mismatches  321;  Indels    0;  Gaps    0;

Qy    185 ttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaactaag 244
          |:| || :|:  ::: :::| |: ::||| |  :||    ||:||| |:    :| |
Db    453 TYTCTCAHTWTMMMMMWWAATWTWWAAAWAAAWTTATWAATWAAAAAAWAWWWWATTTTT 512

Qy    245 ttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctatttgg 304
          :::       : | : :: :| :|      || :||   |   | :|| |:  || :
Db    513 WWWWWTWATTWTTTWAWWTWTAWTAAAAAAAAAWATAATTTAAAWWAATAWATTAAWAAT 572

Qy    305 actgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctcacat 364
            :| |   ||    | | : :   ||         | |      :    |  |
Db    573 TTAWAAWWTATATTAATWTATAAATWTWATTAATATAAAAAAATATTTTTTWATAAAATT 632

Qy    365 cgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgctagt 424
              | | ||:   | |||  |:|:  |:::||     |  | :    :
Db    633 TTTAATAATTTAATTAWTTATTAAATAAWTWATTWWWTAATTAAATAATTTWAAATAWAA 692

Qy    425 atgtgttataaaatcattttaccaattaccttttaaataactgtacgtagtttcatcag 484
          |      | ||||:  :  :|  :|| :    :|||||| :| |  |   |   | |: |
Db    693 AAAAAAAAAAAAAWATWAAWAATWATAWATAAWTTAAAAWAATAAAAWAAWAATWAWWAT 752

Qy    485 tagaactactcacggagctaaaacaaaaaaagttgttctactgataaaagcagagatgat 544
          | |  :|   |      ||:::  ::|:::  | :   :   |:||:|  :   |  |:
Db    753 AATATWWATATATAWTTWTAWWWATWWAWWWTATAWATAWAATAWAAWAWAWATAAATAW 812

Qy    545 gtatgaccgtgaccgtgagctaaagtccaaaaaaaaaaactgctccacaataacgacaaa 604
           ||:       |    :|    ::|| |:|:||    |  |  ||:|  :  ||
Db    813 ATAWATWAAAWAWAWATAWWATWATATAWWAATAWAWAAAAAATWTAATATWAATWATAA 872

Qy    605 acaaagttgtattgtatggcctaaattacagcacactgacaccacacgtatattattctc 664
          : |||  | :|:| ::|    |::|::| |  | : :   :  | |   | |  |   |
Db    873 WAAAAAATAWAWTTWWTWTTTTWWAWWATATAAAWAWATAWAAAWAAAAAAAAAAAAATA 932

Qy    665 tctccattatcacaggatgtaactgtaaaaattttgtatgttaaacatttgtagtaaata 724
              :|:||   :  ||  || | || ::|||:  :|:: :|: | |:|| || : |
Db    933 AWAWWWTWATATTWTTATTAAAWTWTATWWATTWATTWWAWTWTATAWTTWTATWTATAT 992

Qy    725 ttgctagcatttacgtctacggaatttattgaaaaaatgtagtattgttttatataattt 784
          |  || ||: ||  | ||   |::|   |   |:||  || |||  : | :::||: |
Db    993 ATWWTAWTAWATATATTTATTAAWWTATATTTTAWAAAWTAATATATWATAWWWTAWATA 1052

Qy    785 taataaaactgtaaat 800
          ||: ::|| | :  ||
Db   1053 TAWAWWAATTAWTTAT 1068


RESULT   7
AW231748
LOCUS       AW231748      454 bp    mRNA            EST       30-MAR-2000
DEFINITION  687023G08.y2 687 - Early embryo from Delaware Zea mays cDNA, mRNA
            sequence.
ACCESSION   AW231748
VERSION     AW231748.1  GI:6564126
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 687023  row: G  column: 08.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Zea mays"
                     /cultivar="Illinois High Oil"
                     /db_xref="taxon:4577"
                     /clone_lib="687 - Early embryo from Delaware"
                     /tissue_type="embryo"
                     /dev_stage="14, 21, 28, and 35 days after pollination"
                     /lab_host="E. coli SOLR"
                     /note="Organ: embryo; Vector: pBluescript SK; Site_1: XhoI
                     ; Site_2: EcoRI; Library was prepared by Statagene using
                     the Uni-ZAP XR system (Stratagene BN937328-12). Clones
                     were picked by a Q-bot after blue/white selection
                     (ampicillin resistance - use 100 micrograms/microliter).
                     Developed from a pool of equal amounts of RNA from
                     developing embryos sampled at 14, 21, 28 and 35 days after
                     pollination of the Illinois High Oil Maize Strain Cycle
                     90. This closed strain has been selected for high oil
                     concentration for 90 generations and originates from the
                     1890s era open pollinated variety Burr's White"
BASE COUNT       99 a    133 c    154 g     68 t
ORIGIN


  Query Match             2.8%;  Score 58;  DB 10;  Length 454;
  Best Local Similarity   92.4%;  Pred. No. 0.031;
  Matches   61;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   1989 gcggcgctcgcttacctcgcctcgcattccgttcgagcaggggagcggcagtgagaaggg 2048
          || |||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||
Db     16 GCTGCGCTCGCTTAACTCGCTTCGCATTCCGTTCGAGCAGGGGAGCGGCAGTGAGAAGGG 75

Qy   2049 agggaa 2054
          | || |
Db     76 AAGGGA 81


RESULT   8
CNS0167M/c
LOCUS       CNS0167M     1201 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106396
VERSION     AL106396.1  GI:5621701
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
```

project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

```
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"
                     /clone_lib="DrosBAC"
                     /clone="BACN15M24"
                     /note="end : T7"
BASE COUNT      323 a     87 c     79 g    551 t    161 others
ORIGIN
```

```
  Query Match             2.8%;  Score 58;  DB 13;  Length 1201;
  Best Local Similarity   37.0%;  Pred. No. 0.033;
  Matches  211;  Conservative   60;  Mismatches  293;  Indels    7;  Gaps    1;

Qy      39 atcacaatctaaaaatcaggagaaattatatggccatattatagaagcaactaaataaaa 98
           ||   ||| |||||||  |  | |:| |||||    | |      ||  |:  ||  ::|
Db    1184 ATATAAATATAAAAATAAATAWAWAATATATAWNAAAAANTATAAAAAAWAAAATAWWAT 1125

Qy      99 tgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaatacaattct 158
                 | : ::| |  ||   :  ||||| :|| |      || || : : ||:|
Db    1124 ATAAAWAAWAAAWWWWATATTAAATAWAATATAANWAAAAAAAAAAAAAAAAAAWWWTTTHT 1065

Qy     159 tttatacacaacttatatgtgagttctttttctcttgtaactcttattaataaaacattt 218
            |||  :  |: ||| : | :| ||||| | || |   | : :  |: |||| |  |
Db    1064 ANAATATTTWTTTWNTATAWAATWTTTTTTTTTTTTTTTTATAWAWAAAWAAAAAAAAAT 1005

Qy     219 ttggctattaaataatggcaactaagttagcaccactgtaattagattttgtctggaaca 278
           ||    | ||||: |     :  || ||   |  | | | : |:  ||||      || |
Db    1004 TTTAAAAATAAAWTAATTATWAAAATTTTTAAAAATTTTTWATWWTTTTTWAAAAAAAAA 945

Qy     279 atttctctgactaagaagctatttggactgtccttttgccaaacaagtagaaaatggaac 338
           |:     : | :       |||:|  |      :|||    :  ||  | ||||:  ||
Db     944 AWAATATWAAAWTTTTTTTTATWTATAAAWAWAWTTTTTTYWTTAAAAAAAAAAWTTAAA 885

Qy     339 cgctccttaaaaaaccattctcacatcgctgggtgctgaataaaactgaaaacattagct 398
                 | : ||  ||| | | ||           || ||||   ::|| | :   :
Db     884 A-------ATWTAAAAATTATAAAATAAAAWAAAAAAAAAAAAAAAAAAWWAAAAAWTTTW 832

Qy     399 ttttatagctctcgctctctgctagtatgtgttataaaatcattttaccaattacctttt 458
            ||   |  | |           |  || |  | :    | |   ||  || :|  : ::
Db     831 ATTATAAWATTTTTWAAAAAAAAAAAATTTAATTYTTTTTNAAAATAAAAAAWAAAWAWW 772

Qy     459 taaataactgtacgtagtttcatcagtagaactactcacggagctaaaacaaaaaaagtt 518
           :||| || : ||  ||   :    |     |  ||  |   :  | ::| |||||||
Db     771 WAAAAAATWWTAAATATAAWTTAAATNCATAAAACAAAAAAWAATTWHATAAAAAAAAAA 712

Qy     519 gttctactgataaaagcagagatgatgtatgaccgtgaccgtgagctaaagtccaaaaaa 578
                |  | ||||  | |  |  | :|  |::   |::   : | |||  ::||::||
Db     711 AAAAAAAAAAAAAAAAAANATTTTTTTHACAAMMAMMAMMYMMMMCAAAAMAMMAAVMAA 652

Qy     579 aaaaactgctccacaataacgacaaaacaaa 609
           |:| :   :|  |   |||  | | || :||
Db     651 AMACMAMMMTAAAGCNTAAAAAAACAAAMAA 621
```

```
RESULT   9
CNS0021J
LOCUS       CNS0021J     1101 bp     DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL061936
VERSION     AL061936.1  GI:4940214
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR05N11"
                     /note="end : TET3"
BASE COUNT      631 a      7 c     28 g    289 t    146 others
ORIGIN
```

```
  Query Match             2.8%;  Score 57.4;  DB 13;  Length 1101;
  Best Local Similarity   44.1%;  Pred. No. 0.044;
  Matches  130;  Conservative   20;  Mismatches  145;  Indels    0;  Gaps    0;

Qy      30 aaatatccaatcacaatctaaaaatcaggagaaattatatggccatattatagaagcaac 89
           ||| |   ||  | ||   |||||  |  | |||  | |     | |  | | ||  ||
Db     395 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 454

Qy      90 taaataaaatgtgcgttgtattgaaaaaaaaacctatttataacaaacatctgccaagaa 149
            ||| ||||          |  |||||||||   |    | || ||| |       || ||
Db     455 AAAAAAAAAAAAANAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAANAAAAA 514

Qy     150 tacaattcttttatacacaacttatatgtgagttctttttctcttgtaactcttattaat 209
           || ||||  :|| |  :  | |||: | |   || ||||| | |: ||| | | |||  |
Db     515 TATAATTTAWTTTTTTWTTAATTAWTTTTTTTTTTTTTTTTTTWTTAATTTTAATTTTT 574

Qy     210 aaaacatttttggctattaaataatggcaactaagttagcaccactgtaattagattttg 269
           ||:| |  |||    |  :|:|: |      ||| :||  :  |    |:||  |  :
Db     575 AAWAWAAATTTAATAAAAWAWTWTTTAWTTTTAATWTAAWWAAAAAAAAWTTTTAAAWAW 634

Qy     270 tctggaacaatttctctgactaagaagctatttggactgtccttttgccaaacaa 324
           | |       :||   | |  :| ||  :|| :  : | |   |::   ||| :|
Db     635 TWTTTTTTTTWTTAATTAATAWAAAAAAAWATAWAAWTTWTTTATWWATWAAATWA 689
```

```
RESULT  10
CNS00DKY/c
LOCUS       CNS00DKY      928 bp     DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL071865
VERSION     AL071865.1  GI:4948170
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 928)
```

```
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .928
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR27A24"
                     /note="end : T7"
BASE COUNT      262 a     70 c     84 g    321 t    191 others
ORIGIN
```

```
  Query Match              2.7%;  Score 56.4;  DB 13;  Length 928;
  Best Local Similarity   29.9%;  Pred. No. 0.072;
  Matches   86;  Conservative   78;  Mismatches  124;  Indels    0;  Gaps    0;

Qy     566 aaagtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggcc 625
           |||    ||||||||||||       | || ||  | |||| |||     :::  :   :
Db     909 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAWWYTCMAAAMA 850

Qy     626 taaattacagcacactgacaccacacgtatattattctctctccattatcacaggatgta 685
           |:  ::: ::::: ::::::  : ::::::::::: | |   |: |       | |
Db     849 MAMMMCVMMMMMMMMMMMMMMTTMWWHHHWHHHHHHWTTTTTTTWTTTTTTTTTTTTT 790

Qy     686 actgtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctacg 745
           : : :::::|:||| ::: ::::: ::::    : : :: :  ::  :: ||  | ||
Db     789 WWWWWWWWWWWTTTWWWWTWWWWWWWWWWTNNNWNMCHWAWATWWWWWWATATTTWTATW 730

Qy     746 gaatttattgaaaaaatgtagtattgttttatataatttaataaaactgtaaatcgtct 805
           :|: ||:  |||||   |  :|   ||| |:|  | :| :|||| | |  ||  | |
Db     729 TWAWATAWAAAAAAAAAAAAAAWTAAATTTTTWTTTTAWTTWTAAAATTTTTTATTWTTT 670

Qy     806 ggcttcgtttctggatggaggataaatagtgaatacgaatgggaaaca 853
           :|  |:| |  ||  |  |: |: :   |:     ||   |:| |
Db     669 TTTWTTTTWTTTWTATATAAWAWTAWAWAAWAWATATTATATWAWATA 622
```

```
RESULT  11
CNS0091P
LOCUS       CNS0091P      925 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1  GI:4934461
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR19D16"
                     /note="end : TET3"
BASE COUNT      120 a     61 c     61 g    172 t    511 others
ORIGIN
```

```
  Query Match              2.7%;  Score 56;  DB 13;  Length 925;
  Best Local Similarity   19.9%;  Pred. No. 0.087;
  Matches  129;  Conservative  168;  Mismatches  348;  Indels    3;  Gaps    1;

Qy     653 tatattattctctctccattatcacaggatgtaactgtaaaaattttgtatgttaaacat 712
           | | ||||| | | |     |  |         || | |  |||| |   |    | |
Db     280 TNTTTTATTATTTGTNTNNCANNANNTTGNNNAAATNTNTAAATATAAAAAAAAATATTN 339

Qy     713 ttgtagtaaatattgctagcatttacgtctacggaatttattgaaaaaatgtagtattgt 772
           ||  ||| | |  ||  | ||           || ||           |  | |
Db     340 AAATAAAAAAAACTNNTAAAANTTGNNNNNTNTNNATATAGNNNNNNNNNNTNANAANNN 399

Qy     773 tttatataatttttaataaaactgtaaatcgtctggcttcgtttctggatggaggataaat 832
           |||   | |    |  |   |  |||              |      |   |  |||||
Db     400 NTTA---ANTNNNNANGANNATNNAAANAAANAANNNNATNTAAAAAAACAAAAATAAAA 456

Qy     833 agtgaatacgaatgggaaacaccacaacaaccacgccgctgcgttctgcgaatcacatga 892
           | |: |  |||              |     |           | |   ||    | |
Db     457 TATAAWAAATAATANTTTNTNNNNNNANNNNNAATANANNNTTATTTATTAANNNNANNA 516

Qy     893 gcgatcagtgccttgctgttccgtgaacttgcacgcaaggacgagagcctttctgccttt 952
           |          ||    | ::    |:|:      :: :| | ||:    ::|:
Db     517 NANANANANNNNNAGCSMSMCGKKCGSTTBGSTTTTTTSSGSGYGKGCSSGSGBSCSCCS 576

Qy     953 gcatgcaaggacaagagtctttacatgcaaggacaaataactcccacgcgccccaccgtg 1012
           :|  ::   : |      :    |  ::  :: |      | ||| ::::|  : ::: :
Db     577 SCSCSSSSCSCCBCCCCCSCSSYCCSSSBSSSKCSSTSBSCSCCCSSKSVCGTSCSSSSS 636

Qy    1013 ctttggcaagccacatggcaccctgccgatcacaattcacaggcccaggcttccggtggt 1072
           |   :::  ::: :   :  ::: :::  : :   :: :|:|: :::    : :  :
Db     637 CSSSSSSTSSSSTSSSTSKSSSSGSSSSSSSSYTTSKSTSASGSGSWSAGGGSGSTGSTS 696

Qy    1073 cgcgtgccgtgagtctgacaccgcaccacatggccgccgtaggccgtgcctacgcaccaa 1132
           :::: :: :|: : : : :: : :  :  :|:::::| :::: |: :  ::: ::
Db     697 SSSSSSSTSTSSSSVSSGSKSSTBSSGSBSSSGSSSSSSTSSBBSCTSTSSSSSSSYSSS 756

Qy    1133 ggcgactcgtggtgccaggctctcggcggctttggagtcggtgccatgccgcggggtccg 1192
           :    |  :::: ::: :|: |: :    :::||::: ::  | |: |    |||:
Db     757 TCSCCTCCCSYSYSSSTSSSSSTSWGSTSGSSSSSSVGTSSSSDSTSTCCSCCCYMCTCCS 816

Qy    1193 tggaccgctcctaggggccaggacgaagctgcaccgcacaagcgggccgcgcgtgactcc 1252
           | :::|    :  |::::: :| :   || || :::   : | :: |:::::: : : ::
Db     817 TYBMBCYTSTSCGGSSSSSGKGGVTKCGCGGCGSSSTNGMBGTSSACSSSSSSCSSSSVS 876
```

```
Qy   1253 gtgaccgtgaggcgggcgtaaccaggagcttccgccacgcttgagacc 1300
          : : ::  : ::::::|:: : :  ::|:   ::|:: :|:  | | ::
Db    877 SSSKSSASSSSVSSSGSSGVSSNSSSASKSSSSGSVSSGSGSGSGSVS 924


RESULT 12
AU086608
LOCUS       AU086608      500 bp    mRNA            EST       27-JAN-2001
DEFINITION  AU086608 Sugano Malaria cDNA library Plasmodium falciparum cDNA
            clone XPFn3029, mRNA sequence.
ACCESSION   AU086608
VERSION     AU086608.1  GI:12388749
KEYWORDS    EST.
SOURCE      malaria parasite P. falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Watanabe,J., Sasaki,M., Suzuki,Y. and Sugano,S.
  TITLE     FULL-malaria: a database for a full-length enriched cDNA library
            from human malaria parasite, Plasmodium falciparum
  JOURNAL   Nucleic Acids Res. 29 (1), 70-71 (2001)
  MEDLINE   20574754
COMMENT     Contact: Junichi Watanabe
            Institute of Medical Science
            The University of Tokyo, Department of Parasitology
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Tel: 81-3-5449-5378
            Fax: 81-3-5449-5410
            Email: jwatanab@manage.ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .500
                     /organism="Plasmodium falciparum"
                     /strain="3D7"
                     /db_xref="taxon:5833"
                     /clone="XPFn3029"
                     /clone_lib="Sugano Malaria cDNA library"
                     /dev_stage="erythrocytic stage"
BASE COUNT      214 a     23 c     22 g    223 t     18 others
ORIGIN


  Query Match             2.6%;  Score 54.8;  DB 10;  Length 500;
  Best Local Similarity   52.4%;  Pred. No. 0.15;
  Matches  110;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy      1 tctagatttttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggag 60
          | ||  |||| |   |||  |      |   || ||    | | ||| ||| |||    |
Db    124 TATATTTTTTGTAAAAATGTATAAAAAAAGAAAATAGTNTTAAAATNTAATAATTTTATAA 183

Qy     61 aaattatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaa 120
           |||||          |||| ||  ||||    ||||| |   || |||| ||||| |
Db    184 NTATTATNNATAATATATATATAATAAACTTTTCAAAATTTTATTTATATTAAAAAATAT 243

Qy    121 acctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtga 180
          |  ||||  ||| |||     |   || |  | |||  || || |  ||| |   |    |
Db    244 ATATATTATTAAAAAATGAATAAAAATATCATAATNTTTATAAAANCAATTATAANTATA 303

Qy    181 gttctttttctcttgtaactcttattaata 210
            || | | |    | ||| | |||| ||||
Db    304 TTTTTATATAGTGTATAAATATTATGAATA 333


RESULT 13
CNS017Z2
LOCUS       CNS017Z2      869 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37C17 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108680
VERSION     AL108680.1  GI:5628984
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 869)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .869
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"
                     /clone_lib="DrosBAC"
                     /clone="BACN37C17"
                     /note="end : SP6"
BASE COUNT      269 a     94 c     76 g    223 t    207 others
ORIGIN


  Query Match             2.6%;  Score 54.4;  DB 13;  Length 869;
  Best Local Similarity   37.3%;  Pred. No. 0.19;
  Matches   97;  Conservative   47;  Mismatches  116;  Indels    0;  Gaps    0;

Qy    505 aaacaaaaaaaagttgttctactgataaaagcagagatgatgtatgaccgtgaccgtgagc 564
          | | ||||::|    || :|   | |||  |   |: :: ::  |     |     |
Db    606 AGAAAAAAMMAWAAWTTAWAAAAAAAAAAWTAATWAWWWWWWWAAATAAAAAAAAAAAA 665

Qy    565 taaagtccaaaaaaaaaaactgctccacaataacgacaaaacaaagttgtattgtatggc 624
           |||    |||||||||||| :  |||| | ||  | |:|| |||   :|  :|:
Db    666 AAAAAAAAAAAAAAAAAAAWCWTCCAAYAMAAAHAAAWAAAAAAATAAAWAAATWAHATA 725

Qy    625 ctaaattacagcacactgacaccacacgtatattattctctctccattatcacaggatgt 684
          |:|:||: | || :   : :: | |  :: ||:  :|| | |   || :|: :  :| |
Db    726 ATWAWTTWAAACAAWAAAWTWHWATAWAWMAATWTAWCTTTTTTTTTTTTWCHWWWTWTAT 785

Qy    685 aactgtaaaaattttgtatgttaaacatttgtagtaaatattgctagcatttacgtctac 744
          |::   |: :: ||:  :| |  :: :||| |  |    :|| ||   ||::    :|
Db    786 AMMAMAAMCMMMTTYMMWTTTATWWTWTTTTTTTTTTTAWTTTWTATWTTTWWAWAAWAA 845

Qy    745 ggaatttattgaaaaaatgt 764
            ::::| || :||| || |
Db    846 TAWWWWTTTTAWAAATATWT 865


RESULT 14
CNS0091P/c
LOCUS       CNS0091P      925 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1  GI:4934461
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR19D16"
                     /note="end : TET3"
BASE COUNT      120 a     61 c     61 g    172 t    511 others
ORIGIN
```

```
  Query Match             2.6%;  Score 53.8;  DB 13;  Length 925;
  Best Local Similarity   12.5%;  Pred. No. 0.26;
  Matches   41;  Conservative  168;  Mismatches  119;  Indels    1;  Gaps    1;

Qy   1001 cgccccaccgtgctttggcaagccacatggcaccctgccgatcacaattcacaggcccag 1060
          |:::|: :|: :::  ::: :: : :  ::| :|: ::::    ::   : : :: :: :
Db    925 CSBSCSCSCSCSSBSCSSSSMSTSSSNSSBCSSCSSSBSSSSTSSMSSSSBSSSSGSSSS 866

Qy   1061 gcttccggtggtcgcgtgccgtgagtctgacaccgcaccacatggccgccgtaggccgtg 1120
          :: |::   :   ::: |||| |: :|   : :::|  :|:| |: :: :| ||:  | :
Db    865 SSGTSSACVKCNASSSCGCCGCGMABCCMCSSSSSCCGSASARGVKVRASGGAGKRGGGS 806

Qy   1121 cctacgcaccaaggcgactcgtggtgccaggctctc-ggcggctttggagtcggtgccat 1179
             :: :: | ::: : |: : :: : ::: : : :::||    :||: ::: :::
Db    805 GGASASHSSSSACBSSSSSCSASCWSASSSSSASSSRSRSGGGAGGSGASSSRSSSSSSS 746

Qy   1180 gccgcggggtccgtggaccgctcctaggggccaggacgaagctgcaccgcacaagcgggc 1239
          : |:::::: ::: :: :::: |: : :: :| :: :: |:  :: ::::|: |:|:  |
Db    745 ASAGSVVSSASSSSSSCSSSVSCSSVASSMSCSSBSSSSASASSSSSSSSSASCASCSCCC 686

Qy   1240 cgcgcgtgactccgtgaccgtgaggcgggcgtaaccaggagcttccgccacgcttgagac 1299
           : :|: : : :::  :::: : ::|:::  | ::|:: :   ::::  :::  :|  :
Db    685 TSWSCSCSTSASMSAARSSSSSSSSCSSSSMSASSSASSSSASSSSSSGSSSSSGSACGB 626

Qy   1300 cacgtgacggcgcagaggagctccacgcg 1328
          :::: | :|:::  :  : :: ::  :::
Db    625 SMSSGGGSGSVSASSGMSSSVSSSGGRSS 597
```

```
RESULT  15
CNS014PQ/c
LOCUS       CNS014PQ      987 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL104456
VERSION     AL104456.1  GI:5616067
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
```

```
REFERENCE   1  (bases 1 to 987)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .987
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"
                     /clone_lib="DrosBAC"
                     /clone="BACN12P22"
                     /note="end : SP6"
BASE COUNT      257 a    122 c    122 g    241 t    245 others
ORIGIN
```

```
  Query Match             2.6%;  Score 53.4;  DB 13;  Length 987;
  Best Local Similarity   25.6%;  Pred. No. 0.32;
  Matches  117;  Conservative  119;  Mismatches  221;  Indels    0;  Gaps    0;

Qy      3 tagatttttttttcaattcaccccgagtaaatatccaatcacaatctaaaaatcaggagaa 62
            :: ::::|||| : ||          |::| |   :: : ::| | ::::: :    ::
Db    949 WWWWWWWTTTTTTWTTTTTTTTTTTTTAAWWTTTAWTWWWWWHWTTTYHWWWWTWTTTTWW 890

Qy     63 attatatggccatattatagaagcaactaaataaaatgtgcgttgtattgaaaaaaaaac 122
          ||::| |    :|:: ::  :      |::: |:||| |   ::  |   |||::||::
Db    889 ATWWTTTWATTWTWWAWYCYWTATTTWTWWWAAWAATTTATAWWAAAAAAAAAWWAAWMT 830

Qy    123 ctatttataacaaacatctgccaagaatacaattcttttatacacaacttatatgtgagt 182
           ||:: :|| ||  ::    .  |  |||: | |  ||   ||   || : | |  | :
Db    829 NTAWWAWTATTAATTWWWATWAATWAATWAATTAWTTAATTAATWAAWWAAAAATTTWAA 770

Qy    183 tctttttctcttgtaactcttattaataaaacatttttggctattaaataatggcaacta 242
          | |  |  : || |     ::||::  |:|:  |:|||   | |:  | ||    || :|
Db    769 TTTAATAAWTTTCTTTAAAWWATWWTAAWAWTTTWTTTWWTTTTWTTAAAAAAAAAAAWA 710

Qy    243 agttagcaccactgtaattagattttgtctggaacaatttctctgactaagaagctattt 302
          : ::|  |  | : ::|| :   |||  | |       :||| |    ||||       ::|
Db    709 WWWWAAWAWNANWTWWATAWTTATTTTTTTTTTTWTWTTTTTTATTCTAATTTTWGTWWT 650

Qy    303 ggactgtccttttgccaaacaagtagaaaatggaaccgctccttaaaaaaccattctcac 362
          :   : |  ::|::         |  |::| ::: :          |::::::::  : : :
Db    649 KWTTKTTKTKKTKKKGGGGGGGGGGGMVAGWKRTMAAAAAAAMMAMMMMMMMNVMMVYY 590

Qy    363 atcgctgggtgctgaataaaactgaaaacattagctttttatagctctcgctctctgcta 422
          :: ::   :  :   | || : :   :   : : |: : ||:| | : : | |  ::
Db    589 YYYYHYTTYHTTWTTTTCTAAWHMHYCCMYCAWCYATHAWCTAKCACAMAYACACGCMWT 530

Qy    423 gtatgtgttataaaatcattttaccaattacctttttt 459
            |     | |    |::||:: |: :::   ||:|:
Db    529 ARAGTGAATTTTCTTTMWTTHWTCYYHYHTTTTTWTW 493
```

```
Search completed: December 19, 2001, 21:47:56
Job time: 11632 sec
```